WAVELETS, ICA AND STATISTICAL PARAMETRIC MAPPING: WITH APPLICATION TO AGITATION-SEDATION MODELLING, DETECTING CHANGE POINTS & TO NEUROINFORMATICS

A thesis submitted in partial fulfilment of the requirements for the Degree

of Doctor of Philosophy in Statistics

at the University of Canterbury

by In Kang

University of Canterbury

2011

Abstract

The wavelet methods developed, advocated and used in this thesis are primarily based on the discrete wavelet transform (DWT), wavelet thresholding and density estimation via wavelet smoothing. First a suite of wavelet techniques are advocated, based on the DWT, and applied successfully to assess whether an ICU patient's simulated[π] agitation-sedation (A-S) status reflects their true dynamic A-S profile. The use of quantitative modelling to enhance understanding of the A-S system and the provision of an A-S simulation platform are key tools in this area of patient critical care. Secondly novel wavelet density metrics are developed, a wavelet time coverage index (WTCI) and a wavelet probability band (WPB), based on Bayesian density estimation. This led to the development of two numeric metrics, the average normalized wavelet density and the relative average normalized wavelet density; both shown to be in close agreement to our DWT and earlier metrics.[π] The DWT and WPB approaches also yield excellent visual assessment tools and are generalisable to any study involving bivariate time series of a large number of units (patients, households etc) and of significant length.[℘] Wavelet thresholding and independent component analysis (ICA) are tested as denoising methods, and applied to brain image data, as part of the neuro-informatics study of Turner et al. (2003). ICA methods are then implemented for denoising all the cerebral function data, at a voxel by voxel basis. This is performed as a pre-processing step to the creation of statistical parametric maps (SPMs), used to model brain function with respect to personality as non-linear models. The results derived from our novel SPM-ICA approach support the theory of a biological basis for personality and report more de/activation clusters in the brain, as related to specific personality traits than Turner et al. (2003). Our work gives credence to a growing body of thought for the need of non-linear modelling in psychometric research (Cloninger, 2008). Our work also has the potential to increase momentum for patient specific drugs for depressives. Lastly we develop a DWT based methodology for change point analysis that uses modified, maximal overlap DWT (MODWT) coefficients to link to a novel shifting DWT (SDWT) methodology - a combination of the DWT and MODWT for the change point detection problem, which is shown to provide an accurate and computationally efficient change point location method. SWDT can be generalized to find multiple change points, by way of a binary segmentation procedure and iteration.

[π] Metrics and the simulated profiles are based on differential equations and non-parametric regression methods as derived by Chase et al. (2004) and Rudge et al. (2006a, 2006b).

[℘] The A-S time series profiles studied were of significant length (range [3,001-25,261] time points in minutes).

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible. Early in the process of completing a Ph.D, it became quite clear to me that a student cannot complete a Ph.D dissertation alone. I would like to gratefully acknowledge my supervisory team at the University of Canterbury (UC), Dr Marco Reale and Dr Carl Scarrott and my two off site supervisors, Professor Irene Hudson, from the School of Mathematical and Physical Sciences, the University of Newcastle, and Dr Robin Turner from the School of Public Health, the University of Sydney. Without their invaluable suggestions, ideas, warm encouragement and thoughtful guidance during this study, I would have been lost. They contributed to the ideas, theory and applications presented and to my understanding gained.

My journey towards this thesis spans several years of work in the wavelets and the neuroinformatics domain. I owe my most gratitude to my supervisor Professor Irene Hudson particularly for mathematical guidance, theory and adding rigour to chapters 1 to 5 inclusive in my thesis. She was untiring in her help during my difficult moments and despite working in Australia, Professor Hudson continued to supervise off-site from 2006 to now. Her understanding, encouraging and personal guidance have provided a good basis for the present thesis. She also helped me seek valuable momentum and advice early on from wavelet gurus - Professor Don Percival from the University of Washington and Professor Andrew Walden from the Imperial College of Science, Technology and Medicine, London.

I thank Professor Geoff Chase and Dr Andrew Rudge of the Department of Mechanical Engineering and Centre for Bioengineering, UC, for access to the agitation sedation data which is integral to the application in chapters 2 and 3. I am indebted also to Dr Robin Turner for allowing me to access her unique brain function and psychometric data (results in Chapters 4 and 5). I also thank SPM for allowing me to use all materials available on the Statistical Parametric Mapping (SPM) website (http://www.fil.ion.ucl.ac.uk/spm/) in this thesis. Thanks also to Professor Stephen Roberts from Oxford University who gave advice on and permission to use his hierarchy of ICA models, and to Mukesh Motwani from the University Nevada for access to his diagram of Classification of image denoising methods. I also thank Dr Michael Hermes, Department of Psychology, University of Heidelberg, Heidelberg, Germany for kindly allowing me to use his personality table.

I am most grateful to my friends, Jiyon, kyung-A, Zita, and my soul mate Sun-Hee for being my surrogate family during the years that I have studied and for their continued moral support.

In particular, I wish also to express my gratitude to my UC office mates, Dr Jean Gong and Wen Eng Ong for their continued encouragement and suggestions during this work. I would also like to thank Steve Gourdie and Paul Brouwers of the School of Mathematics and Statistics, UC for their untiring assistance and expertise with all types of computer technical problems - at all times.

Finally, many thanks to my family for the help and support they provided me throughout the thesis process. My husband Min-Gyu Kim, your patience, love and encouragement have upheld me, during many study years. The encouragement and support from our handsome and smart two sons Han-young Kim and Han-Jun Kim was and is a powerful source of inspiration and energy. I am forever indebted to my family for their understanding, endless patience and encouragement when it was most required. It is, however, not possible to list all here. I greatly appreciated their support.

Publications Related to the Thesis

Book Chapter:

Kang, I., Hudson, I.L., Rudge, A., Chase, G. (**2012**). A book chapter proposal submitted and entitled "*Density estimation and wavelet thresholding via Bayesian methods: A Wavelet Probability Band and related metrics to assess agitation and sedation in ICU patients*" In: Discrete Wavelet Transforms, InTech publishers, http://www.intechweb.org, Vienna, Austria. ISBN 980-953-307-580-3.

Kang, I., Hudson, I.L., Rudge, A., Chase, G. (**2011**). "Wavelet signatures and diagnostics for the assessment of ICU Agitation-Sedation protocols", for book entitled *Discrete Wavelet Transforms-Biomedical Applications*, InTech publishers, http://www.intechweb.org, Vienna, Austria, ISBN: 978-953-307-654.

Refereed Conference Paper:

Kang, I., Hudson, I.L., Rudge, A., Chase, G. (**2005**) "Wavelet Signatures of Agitation and Sedation Profiles of ICU patients", International Workshop on Statistical Modelling, IWSM 2005, Francis, A.R., Matawie, K.M., Oshlack, A. and Smyth, G.K. (Eds). Sydney 10-15 July, pp 293-296.

Refereed journal papers in preparation:

Kang, I., Hudson, I.L., Rudge, A., Chase, G. "Density estimation and wavelet thresholding via Bayesian methods: application to ICU data", In prep for *Computer Methods and Programs in Biomedicine.*

Kang, I., Hudson, I.L. "Independent Component Analysis (ICA) of brain images with respect to ICA constructs of personality and temperament," In prep for *NeuroImage*.

Kang, I., Reale, M., Scarrott, C.J. "Detecting a change point for time series using wavelets," In prep for *Computer Methods and Programs in Biomedicine*.

Conference Proceedings:

Kang, I., Hudson, I.L., Turner, R., Reale, M., Scarrott, C.J. (**2007**) "ICA and SPM of the relationship between personality and brain blood flow in normal males," New Zealand Statistical Association conference , July 4-6, 2007, Christchurch, New Zealand.

Kang, I,, Hudson, I.L and Turner, R., (**2006**) "Independent Component Analysis (ICA) of brain images with respect to ICA constructs of personality, character and depression: Is there a link?" Statistical Society of Australia and the New Zealand Statistical Association (ASC/NZSA) Conference, 3-6 July 2006, Auckland, New Zealand.

Hudson, I.L., **Kang**, I., Rudge, A.D., Chase, J.G., Shaw, G.M. (**2004**) "Wavelet signatures of agitation and sedation profiles: a preliminary study," The New Zealand Physics & Engineering in Medicine (NZPEM) conference, Christchurch NZ Nov 22-23, 2004.

List of Abbreviations

ACC	Anterior cingulate cortex
AND	Average normalized density
ART	Atheoretical Regression Trees
AWGN	Additive White Gaussian Noise
BP	Bai and Perron
BSS	Blind source separation
C	Cooperativeness
CFT	Creative Functioning Test
CWT	Continuous wavelet transform
Daub4	Daubechies wavelet filter of length 4
DWT	Discrete wavelet transform
EEG	Electroencephalography
EPI	Eysenck Personality Inventory
EPQ	Eysenck Personality Questionnaire
FA	Factor analysis
FGn	Fractional Gaussian noise
fMRI	Functional Magnetic Resonance Imaging
FT	Fourier transform
FTAW	First Tree-Adapted Wavelet Shrinkage
GDD	Generalized Gaussian distribution
GLM	General Linear Model
HA	Harm avoidance
HIV	Human immunodeficiency virus
ICA	Independent Component Analysis
ICI	Interpolated confidence interval
ICU	Intensive care unit
IDWT	Inverse discrete wavelet transform
KSP	Karolinska Scales of Personality
LaDaub (8)	Least asymmetric Daubechies wavelet filter
LM	Linear model
LRD	Long-range dependent
MAP	Maximum A Posterior
MCC	Medial cingulate cortex

MEG	Magnetoencephalography
MJO	Madden-Julian Oscillation
MMSE	Minimum mean square error
MNI	Montreal neurological institute
MODWT	Maximal overlap discrete wavelet
MRA	Multiresolution analysis
MRI	Magnetic resonance image
MSE	Mean square error
NEO-FFI	Neuroticism-Extroversion-Openness Five-Factor Inventory
NS	Novelty seeking
P	Persistence
PA	Pyramid algorithm
PANAS	Positive and Negative Affect Scales
PCA	Principal Component analysis
PD	Pharmacodynamic
PDF	Probability density function
PDWT	Partial discrete wavelet transform
PET	Positron Emission Tomography
PK	Pharmacokinetic
RAND	Relative average normalized density
rCBF	Regional cerebral blood flow
RD	Reward dependence
EPQ-R	Revised Eysenck Personality Questionnaire
RFT	Random Field Theory
SD	Self directedness
SDWT	Shifting discrete wavelet transform
SPECT	Single Photon Emission Computerized Tomography
SPM	Statistical parametric mapping
ST	Self transcendence
STPI	State-Trait Personality Inventory
STTAI	Spielberger state-trait anxiety inventory
TCI	Temperament and Character Index
TCI-R	Revised Temperament and Character Inventory
TPQ	Tridimensional Personality Questionnaire
TI	Tracking index
WCORR	Wavelet correlation
WCCORR	Wavelet cross-correlation

WCCOVA	Wavelet cross-covariance
WPB	Wavelet probability band
WTCI	Wavelet time coverage index

Table of Contents

List of Tables

List of Figures

Chapter 1

1 Introduction

1.1 Literature on Wavelets: Brief History and Applications

1.1.1 History of wavelets

Wavelets have emerged as powerful mathematical tools for the analysis of complex datasets (Gencay et al., 2002; Hudson et al., 2010; Hudson et al., 2011). A wavelet is a so-called little wave, or a brief wave, which as distinct to sine or cosine waves, extend infinitely with a particular frequency and phase. Wavelets are finitely extended or compactly supported; with their oscillations decaying more or less rapidly to zero. After Fourier, the first orthonormal basis was constructed by Alfred Haar around 1910 and was subsequently extended to time-frequency analysis by Dennis Gabor and John von Neumann in the late 1940s. The work of Jean Morlet and Alex Grossman in the 1980s lead to the notion and term "wavelet". The construction of a family of orthonormal bases (with compact support, arbitrary regularity or an arbitrary number of vanishing moments) by Ingrid Daubechies (1992) paved the way for numerous wavelet applications in the area of signal processing. As Bullmore et al., (2004) also states the wavelet analysis can be performed by decomposing variance or energy of a time series by an orthonormal basis of wavelets, its weighted coefficients that are associated with the amount of energy in a time series at a particular scale and time.

A review of the historical development of wavelets is provided by Jaffard et al. (2001). A detailed exposition of wavelets can be found in Burrus et al. (1998). Recent and key works on wavelets are by: Mallat & Peyré (2009); Percival & Walden (2000); Gencay et al. (2002); Ogden (1997) and Vidakovic (1999). Most recent wavelet method with application using R was performed by Nason (2008), see also the exposition was given by Jansen (2001) and Silverman & Vassilicos (2000).

Statistical issues in wavelet analysis of time series are addressed in detail by Percival and Walden (2000). Earlier overviews of wavelet applications in biomedical image processing, as well as early seminal work on fMRI, are provided by Aldroubi and Unser (1996), and Laine (2000). More recently Bruce and Gao (1996) describe implementation of wavelet methods in S-PLUS. Software platforms used in this thesis include: Matlab's Wavelet toolbox (Misiti et al, 1997), the R

package's waveslim (Whitcher, Whitcher, & Gpl, 2006), WaveThresh (Nason, 2008), and strucchange (Zeileis et al., 2002).

1.1.2 Wavelet applications

Since the introduction of wavelets in the geophysical literature by Goupillaud et al. (1984), the application of wavelets to time series (Percival & Walden 2000; Kang et al., 2005; Hudson et al., 2010a; Hudson et al., 2010d) and spatial and image data (Whitcher et al., 2005; Bullmore et al., 2003, 2004) has increased significantly. Wavelets are well suited to signal processing, particularly the analysis of biological signals and images (e.g. human brain imaging and Electroencephalography (EEG) data) which often possess fractal or scale invariant properties (Bullmore et al., 2003). Wavelets allow the decomposition of a time series with respect to two independent variables, namely, time and scale, and also the decomposition of an image with respect to location and extent. Indeed wavelets have wide application in many areas: e.g., signal processing (Percival et al., 2001); climate (Hudson et al., 2010a, Hudson et al., 2010b); environmetrics (Percival et al., 2007; Craigmile et al., 2002); atmospheric research (Whitcher, et al., 2000; Percival et al., 2005; Overland et al., 2004); oceanography (Rothrock et al., 2008; Morabito et al., 2008; Percival et al., 2008; Eriksson et al., 2007), geosciences (Whitcher & Jensen, 2000); turbulence (Nichols-Pagel et al., 2008) and in medical applications (Whitcher, 2005a; Whitcher, 2005b).

Some key papers that have applied wavelets in vegetation ecology and related fields include Bradshaw and Spies (1992), Dale (1999), Lark and Webster (1999), Katul et al. (2001) and Csillag and Kabos (2002). Dale et al. (2002) proposed and used the orthonormal wavelet transformation (OWT) in the first study to explore the measured wavelet spectra and co-spectra of land surface fluxes from fractions of seconds to three years, in a multi-scale analysis of vegetation surface fluxes. Katul et al. (2001) showed that the OWT provides a robust framework for analyzing the (co-) spectral properties of long-term flux records, which manifest both frequency shifts in time, and multiple gaps or missing data. Such features are common to much ecological, geophysical, environmental and phenological time series data. Via wavelets Katul et al. (2001) showed that three broad categories of time scale should be considered when describing the temporal dynamics of land surface fluxes: namely, turbulent time scales (seconds to an hour), meteorological time scales (hours to days), and seasonal time scales (weeks to year).

In numerical analysis and functional analysis, a discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. Whitcher et al. (2000) were the first to propose a multi-scale analysis of covariance between two time series using the DWT.

In atmospheric science for example, understanding the bivariate relationship between two time series is vital. Consider, by way of example, the study of Whitcher et al. (2000) who identified the Madden-Julian Oscillation (MJO) using bivariate spectral analysis; between the station pressure and zonal wind components at Canton Island specifically the co-spectrum and magnitude squared coherence. Whitcher et al's (2000) novel technique for bivariate Gaussian time series utilized the maximal overlap DWT (MODWT), which may be considered to be a non-decimated version of the orthonormal DWT, important in the analysis of geophysical processes (Percival & Guttorp 1994; Percival & Mofjeld 1997).

Furthermore, Whitcher et al. (2000) defined the wavelet correlation (denoted in this thesis by WCORR) and wavelet cross-correlation (denoted in this thesis by WCCORR) statistics. The WCORR and WCCORR statistics were integral to their analysis of El-Nino (Southern Oscillation events and the Madden-Julian Oscillation (MJO)) using a 35+ year record thus demonstrating this method to be an excellent alternative to traditional (Fourier) cross-spectrum analysis. It was shown how the wavelet cross-correlation between ENSO events and the MJO both quantifies and also visually displays how the association between the two processes changes with scale (Whitcher et al., 2000).

A similar approach was taken in a recent a study relating the flowering of four Eucalypt species with climate by Hudson et al. (2010a, b). Further details on these two studies are given in the next section. Likewise an analogous approach is taken in Chapter 2 of this thesis in an analysis of agitation and sedation time series of intensive care unit (ICU) patients. In Chapter 2 wavelet based WCORR and WCCORR statistics, is demonstrated to be useful diagnostics in determining the accuracy of the simulated (modelled) agitation sedation profiles in reflecting a patient's true recorded agitation sedation profile (time series of a significant length) in intensive care unit (see also the early work of Kang et al., 2005).

Potentially complex patterns of cross-correlation are easily decomposed using the wavelet cross-correlation (WCCORR) on a scale by scale basis, where each wavelet cross-correlation series is associated with a *specific physical time scale* (Whitcher et al., 2000; Katul et al., 2001; Percival et al., 2004). This scale-by-scale interpretation is possible as the energy in a time series is preserved in its DWT coefficients (Whitcher et al., 2000; Percival & Walden 2000; Bullmore et al., 2003). Percival et al. (2004) illustrated the use of the DWT in a study of vegetation coverage in the Arctic region. Percival et al. (2004) showed how, given the DWT coefficients, the vegetation time series could be reconstructed perfectly by a multi-resolution analysis (MRA). MRA re-expresses each time series as the sum of a new set of time series (called details and a smooth), each of

which is associated with variations at a particular scale. It was shown how MRA allowed for the identification of certain decades (in the series for the boreal group) for which the year-to-year variations were smaller than usual (Percival et al., 2004).

Phenology is the study of the timing of recurring biological events (e.g. first flowering, first bird arrival). The cause of this timing with regard to climate has had a long history with agricultural phenological calendars (Hudson & Keatley, 2010; Keatley & Hudson, 2010; Hudson, 2010; Sparks & Tryjanowski, 2005; Fitter & Fitter, 2002). Changes in phenological processes have huge consequences to human health, biodiversity, forestry, agriculture; thus enhancing phenology's value as a generic indicator for changes in human and ecological systems. Of particular focus in this introduction is the recent Australian phenological work of Hudson (2010) and Hudson et al. (2010a, 2010b, 2010d, 2011), which as summarised wavelet analysis offer us ways to: [1] identify spatial and climate niche across species; [2] decompose time series into its sub components (e.g. trends, oscillatory modes or seasonality, change-points and noise); [3] establish whether a given species is uniquely influenced by climate through the year (i.e. has its own climatic signature); [4] determine the relationship between multiple climate indicators; [5] succinctly display how the association between the two processes, say climate and flowering, change with scale and time; [6] identify change points in the series; and [7] identify the primary climatic drivers of flowering or of any phenophase.

Recently wavelets were applied to phenological (flowering) time series data, which are frequently non-stationary and highly noisy (see Hudson et al., 2010). It was demonstrated that wavelets (in this application) were capable of establishingspecific and interpretable (phenologically and biologically) types of cycling and changing, dynamic relationships between climate and at both the annual (and across years) basis.

1.1.3 Wavelets and Brain Imaging

Wavelets give an orthonormal basis for MRA and decorrelation or 'whitening' of non-stationary time series and spatial processes. Also the DWT is used extensively for MRA and in "whitening" of such nonstationary time series (as is the nature of the ICU data discussed and analysed in Chapter 2) and in the analysis of 2D or 3D spatial processes (as is the brain image date of Chapter 5). Whitening of an autocorrelated time series, performed by taking its wavelet transform can assist resampling or efficient linear model parameter estimation. The wavelet transform has also proven to be a useful basis for nonparametric regression (see Chapter 3), denoising (see Chapter 4) or compression. Because of their underlying fractal or scale-invariant properties, biological

signals and image (eg. human brain imaging data) are suitable for wavelet analysis. Wornell (1993, 1999) provides a comprehensive argument for the general optimality of wavelet representations for the analysis of fractal signals.

Bullmore et al. (2003) recently defined some key properties of the DWT and reviewed its applications to the statistical analysis of functional magnetic resonance imaging (fMRI) data. fMRI data invariably exhibits nonstationary features at several scales. For the analysis of time series, Bullmore et al. (2003) discuss issues on resampling by 'wavestrapping' of the wavelet coefficients, methods for efficient linear model estimation in the wavelet domain, and wavelet-based methods for multiple hypothesis testing, all of which are somewhat simplified by the decorrelating property of the DWT (Percival & Walden, 2000). Specifically Bullmore et al. (2003) focus on dyadic, orthonormal wavelets (the focus of chapters 2-4), because related methods for time-invariant or undecimated wavelet transformations, result in a redundant or nonorthogonal multiresolutional decomposition of time series (see also Percival & Walden, 2000 for details).

Note that in this thesis modelling brain activations or deactivations, with respect to psychometric constructs of personality and character (Cloninger, 1994, 2008) is achieved using SPM (Friston et al., 2007) not by wavelet analysis. Such work is the topic of future work. Further details on SPM are given below and in Chapter 5.

1.2 Literature on SPM and ICA: Neuro-informatics

1.2.1 Independent component analysis (ICA)

Independent component analysis (ICA) (Hyvärinen et al., 2001b; Stone, 2005; Lee, 1998) is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals (see also Hyvärinen & Oja, 2000; Hyvärinen, 2001a; Lewicki & Olshausen, 1998a; Lewicki & Sejnowski, 1998b; Roberts & Everson, 2001; Stone, 2004). Specifically ICA aims to find underlying factors or components from multivariate (multidimensional) statistical data. What distinguishes ICA from other methods is that it looks for components that are both *statistically independent*, and *nongaussian* (Hyvärinen et al., 2001b).

ICA defines a generative model for the observed multivariate data (Hyvärinen et al., 2009), where the data is typically given as a large database of samples. In the model, the data (variables) are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are called the independent components (ICs) of the observed data, and are assumed to be both nongaussian and mutually independent. These independent

components, also called sources or factors, can be found by ICA. ICA is superficially related to principal component analysis (PCA) and factor analysis (FA) (Hyvärinen et al., 2001b).

Given its generality the ICA model is applicable to many different areas, including spatio temporal images, document databases, in the derivation of economic indicators, in econometrics and finacial mathematic and in the creation of psychometric indices and constructs (Hyvärinen et al., 2009; Hyvärinen et al., 2001b). Specifically in brain imaging, one often has different sources in the brain emitting signals that are mixed up in the sensors outside of the head, as is the case in the basic blind source separation (BSS) model (Acharyya, 2008). In many cases, the measurements inputted into ICA are given as a set of parallel signals or time series; the term blind source separation is used to characterize this problem. In econometrics, one often has parallel time series, and ICA is able to decompose these into independent components thus giving insight into the structure of the data. A somewhat different application is that of image feature extraction, where one wants to find features that are as independent of each other as possible.
In order to simplify and reduce the complexity of the problem, typically ICA algorithms utilise the following tools; centering, whitening and dimensionality reduction, as preprocessing steps. Both whitening and dimension reduction are accomplished via the implementation of PCA or by singular value decomposition (SVD) (Hyvärinen et al., 2001b). Whitening is a process that guarantees that all dimensions are treated equally, *a priori*, before running the algorithm. Algorithms for ICA comprise infomax, FastICA (see Hyvärinen, 1998b; Bingham & Hyvärinen, 2000) and JADE, but this list is by no means exhaustive.

ICA and other types of blind source separation methods have been effectively applied to the problem of separation of spatially independent sources, for the case when fMRI data is measured in a resting state (Beckmann et al., 2005; van de Ven et al., 2004) or during natural stimulation (Bartels & Zeki, 2004). However, the use of ICA to spontaneous electroencephalography (EEG) or magnetoencephalography (MEG) is not straightforward. In general ICA is very good at finding artifacts (Jung et al., 2000a; Vig´ario et al., 2000), but less so in uncovering the components associated with brain (excitation or deactivation) activity. However, recently Hyvärinen et al. (2009) have suggested ICA as a method to model complex and spontaneous EEG/MEG data. Their approach entails a complex mixing matrix used to model sources that are both spatially extended and of differing phases in the diverse EEG/MEG channels. The analysis then implements the complex-valued version of FastICA which uses a robust non-Gaussianity measure (Bingham & Hyvärinen, 2000).

Further details on the procedural steps and mathematics of ICA and on FastICA are given in Chapter 5. ICA is used in this thesis principally to denoise the brain image data, prior to

modelling brain activations or deactivations, with respect to the psychometric constructs of personality and character (Cloninger, 1994, 2003). This work extends the earlier study of Turner et al. (2003) – we incorporate ICA with a non-linear adaptation of SPM (see Chapter 5).

1.2.2 Statistical parametric mapping (SPM)

SPM refers to the creation and evaluation of spatially extended statistical processes, which are used to test specific hypotheses about functional imaging data (Friston et al., 2007; Frackowiak et al., 1997; Frackowiak et al., 2001). SPM is generally employed to identify regionally specific effects (e.g. brain activations) in neuroimaging data, with a view to characterise both functional anatomy and say, disease-related cerebral function and structure changes. Statistical parametric mapping applies Random Field Theory (RFT) for inference pertaining to the topological features of statistical processes that are continuous functions in both time and space. SPMs are thus considered to be continuous statistical processes, in reference to the probabilistic behaviour of random fields (Adler, 2009; Worsley et al., 1992; Friston et al., 1994; Worsley et al., 1995a). Random fields model both the univariate probabilistic characteristics of an SPM and any non-stationary spatial covariance structure. So-called 'unlikely' topological features of the SPM, such as peaks or clusters, are then interpreted as regionally specific effects, which are attributable to the experimental manipulation, if given, or to a specific drug administered, if taken. Clearly the General Linear Model (GLM) is used to explain continuous (image) data by precisely the same means as in conventional analyses of discrete data. RFT thus can resolve the multiple-comparison problem when making inferences over the volume analysed, in that RFT provides a method for adjusting *p*-values for the search volume; playing the same role for SPMs, as does the Bonferroni correction when applied to the situation of discrete statistical tests.

Key stages of analysing imaging data sequences (or time series) in computational neuroanatomy are modeling and inference; with the aim of detecting, and making inferences about, regionally specific effects in the brain. One area not detailed in the literature overview of this Chapter is that area which addresses the integration and interactions among such brain regions, as accomplished via analyses of *functional and effective connectivity* (see NIH Human Brain Project for neuroinformatics research (http://www-bmu.psychiatry.cam.ac.uk/projects/nih-wavelets/, accessed 5th April, 2010). The brain seems to hold fast to two fundamental principles of functional organisation, namely that

of *functional integration* and of *functional specialization,* where the integration within and among specialised areas is mediated by effective connectivity (Friston et al., 1995b).

Brain mapping data are usually analysed via some form of SPM, which, as mentioned above, entails the construction of continuous statistical processes by which to test hypotheses about regionally specific effects (Friston et al., 1991a; Friston et al., 2007). SPMs are images or fields with values that are, under the null hypothesis, distributed according to a known assumed probability density function, usually the Student's t or F-distributions. These are branded as so-called t- or F-maps. The success of SPMs is due primarily to the fact that each and every voxel (*i.e.*, image volume element) are analysed using some standard (univariate) statistical test, usually based on a GLM of the data. The resultant statistics are assembled then into an image - the so-called SPM. It is noteworthy that Turner et al. (2003) were the first to use a novel quartile based adaptation of the GLM in SPM to allow for non-linear trends (in their case between temperament and character traits) and regions of brain activation or deactivation.

Inverting generative models of the data is integral to statistical analysis of imaging data.Inferences are subsequently made using statistics that assess the significance of effects. There are diverse ways to analyse neuroimaging time-series (*e.g.*, see PET, fMRI and EEG and SPECT) as in Chapter 5.

Characterising a regionally specific effect depends on statistical estimation and inference. Inferences in neuroimaging usually involve differences attained by comparison of one group of subjects versus another or, by making comparisons within subjects, i.e. observing and modelling changes over a subject's sequence (time series) of observations. Differences may also be based on structural or anatomical differences (e.g. in voxel-based morphometry (Ashburner & Friston, 2000) or on neurophysiological indices of brain functions (e.g. fMRI). The principles of data analysis are highly similar for all such comparative scenarios. The important issue in such data analysis is the relationship between the neurobiological hypothesis one conjectures and the statistical models adopted and used to test that hypothesis.

The *Statistical Parametric Mapping* approach is voxel based with procedural steps as follows (see also the schema in Figure 1.1):

- Images are realigned (Klein et al., 2009; Ashburner, 2007; Friston et al., 1995), images are then spatially normalised into a standard space (Crinion et al., 2007; Salmond et al., 2002; Ashburner & Friston, 1999), and then images are smoothed (Kiebel & Friston, 2002; Kiebel et al., 1999; Poline et al., 1995).

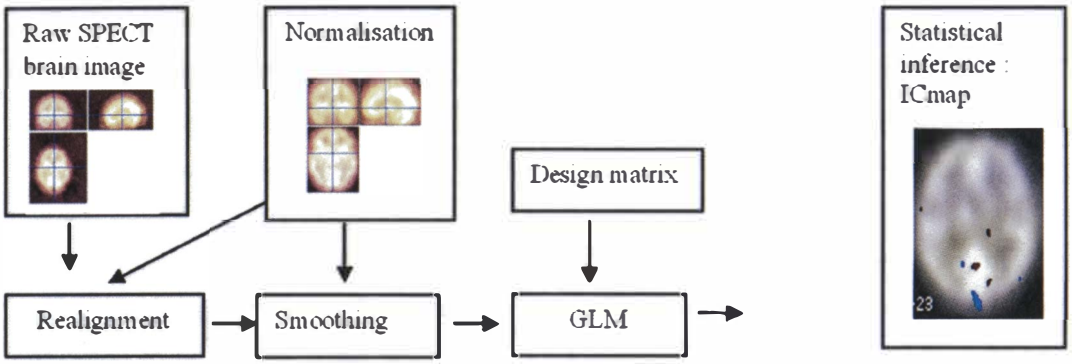


Figure 1.1 Procedural steps from image time series to a final statistical parametric map (SPM): and related inference via the GLM. See http://www.scholarpedia.org/article/Statistical_parametric_mapping_(SPM).

- At each voxel parametric statistical models are assumed, using the GLM to depict the data in terms of both experimental and confounding effects, and of residual variability (Penny et al., 2007; Friston et al., (2005); Penny & Friston (2003); Friston et al., 1995b). For fMRI the GLM is used in combination with a temporal convolution model. Test hypotheses are expressed in terms of GLM parameters. This procedure uses an image whose voxel values can be statistics, a *Statistic Image*, or a *Statistical Parametric Map* (denoted by SPM [t], SPM [Z], SPM [F]).

- The multiple comparisons problem in SPM analysis is addressed using continuous RFT (Hayasaka et al., 2004; Brett et al., 2003; Worsley, 2003; Friston et al., 1991b), whereby the statistic image is assumed to be a good lattice representation of an underlying continuous stationary random field. Inference is then based on corrected *p*-values.

- Bayesian inference can be employed instead of classical inference and this results in so-called Posterior Probability Maps (PPM) (Flandin & Penny, 2007; Penny et al., (2005); Penny & Friston, 2003).

1.3 Wavelet thresholding

Wavelet methods have been shown to be successful in terms of function estimation via term-by-term thresholding of the empirical wavelet coefficients. Many researchers over the last two decades or so have suggested denoising data by thresholding of the wavelet coefficients. The wavelet shrinkage method consists of the following steps: [1] perform a DWT of the original data; [2] threshold the detail wavelet coefficients; and [3] inverse transform the thresholded coefficients - to acquire the actual, original denoised data (Daubechies, 1992; Ogden, 1997; Vidakovic, 1999). The underlying idea is that, wavelet coefficients of small absolute value encode primarily noise and very fine details of the underlying signal. In contrast, important information is considered to be encoded by the coefficients with large absolute value. Removing the coefficients with small absolute value and reconstructing the signal then generate a signal (or image) with reduced noise.

Indeed, when the underlying signal has a sparse wavelet representation, wavelet shrinkage and thresholding methods comprise a commanding method of signal denoising. They are also computationally fast, and automatically adjust to the smoothness of the signal under investigation (see the thresholding applications in Chapters 3, 4 and 6 of the thesis).

Nearly minimax properties for simple threshold estimators, over a large class of function spaces and for a wide range of loss functions, were established by Donoho and Johnstone (see Donoho, 1995; Donoho & Jonstone 1994, 1995 a, b). The notion underlying these wavelet methods is that the unknown function can be well approximated via a relatively small proportion of non-zero wavelet coefficients. Whilst thresholding is a non-linear technique, it is very straightforward because it considers one wavelet coefficient at a time.

Alternative approaches to non-linear wavelet-based denoising can be found in, for example, Abramovich & Silverman, 1998; Charnbolle, De Vore, Lee, & Lucier, 1998; Chipman, Kolaczyk, & McCulloch, 1997; Clyde, Parmigiani, & Vidakovic, 1998; Crouse, Nowak, & Baraniuk, 1998; DeVore & Lucier, 1992; Jansen, Malfait, & Bultheel, 1997; Johnstone & Silverman, 1997; Nason, 1995; Ruggeri & Vidakovic, 1998; Saito, 1994; Simoncelli & Adelson, 1996; Vidakovic, 1998; Wang, 1996, and references therein. We shall next discuss some earlier wavelet thresholding papers, and then move to later papers with specific application to denoising image data, as is performed on the brain image data in Chapter 4, see also Chapter 5 of the thesis.

Donoho and Johnstone (1994) introduced two methods of *global* thresholding: a universal threshold $\lambda = \sqrt{2 \log n}$ and a *minimax* threshold. These techniques apply the common threshold only to wavelet coefficients at higher levels. The set of wavelet coefficients are as one would expect indexed by resolution, and the coefficients are considered one level at a time. See Johnstone and Silverman (1997) and Donoho and Johnstone (1994) for further details. These level-specific thresholds are determined from the coefficients themselves. Donoho and Johnstone (1994) illustrate a thresholding method based on unbiased risk estimation as developed by Stein (1981). Ogden and Parzen (1996a,b) described an alternative thresholding scheme, where the raw wavelet coefficients are examined at each level, and tested to determine if there is sufficient evidence for a significant signal, using standard statistical testing procedures. The largest coefficient is removed if the signal is significant, and then the remaining subset is re-tested. By following this recursive procedure the wavelet coefficients at the current level can be separated, in due course, into a set of coefficients deemed to contain a significant signal, versus a set of coefficients considered to be that of pure noise. The threshold at that level is set, so as to shrink all the 'noise' coefficients to zero - hence the terminology "shrinkage". An intrinsic benefit of this

approach is that the level of smoothing can be controlled by adjusting the significance level of the hypothesis tests, with a smaller (larger) choice of α underfitting (or overfitting) the data. Abramovich and Benjamini (1995) described another multiple hypothesis testing approach to data adaptive wavelet thresholding, which are based on the methods of false discovery rate, as developed by Benjamini and Hochberg (1995).

There are several methods available to calculate and apply thresholds. The simplest threshold is the well-known *Universal* (Donoho & Johnstone 1994) $\sigma\sqrt{2\log n}$, where n is the sample size, and σ^2 is the noise variance. Threshold selection alternatives, founded on minimizing certain optimisation criteria, include the *minimax* (Donoho & Johnstone, 1994), and the *SURE* (Donoho & Johnstone, 1995a) techniques. According to Ogden (1997) thresholds can also be based on *hypothesis testing*, *cross-validation*, and *Bayesian estimation* approaches. The most flexible of the threshold methods, the *Top* method, involves selecting the threshold as a quantile of the empirical distribution of the set of wavelet coefficients. By systematically changing the quantile values, the best threshold for a given application can be found (Ogden 1997). If unknown, the noise variance σ^2 can be estimated, say by the traditional estimates of the sample standard deviation, also by various L_p norms, and by the median absolute deviation (MAD). Level-independent estimates of σ^2 (i.e. one common estimate for all the multiresolution levels in the wavelet decomposition) can be obtained by either including all the detail coefficients, or by using detail coefficients only from a given multiresolution level in the chosen estimate formula. Level-dependent estimates can be obtained by plugging only the detail coefficients corresponding to the given level into the variance-estimating formula.

Threshold, or shrinkage, application functions comprise the *hard*, the *soft*, and the *semisoft* functions (Bruce & Gao, 1995). The hard thresholding function is "keep or kill "operation: coefficients whose absolute values are below a positive threshold will be put to zero, while the other wavelet coefficients are not changed. The soft function is similar to the hard, except that it either shrinks or kills: the kept coefficients are changed by shrinking them towards zero. The semisoft function generalises the hard and the soft functions by utilising two thresholds (Gao & Bruce, 1995). *VisuShrink* (Donoho & Johnstone, 1994) refers to the combination of soft shrinkage with the universal threshold; soft with the minimax threshold is termed *RiskShrink* (Donoho & Johnstone, 1994), and soft together with the SURE threshold is refereed to as *SureShrink* (Donoho & Johnstone, 1995a). A more current advance, such as *BayesShrink* (Chang et al., 2000 a, b), which advocates soft shrinkage in a certain Bayesian framework, claims to outperform *SureShrink* estimates, when applied to denoise images. See Mahmood et al. (2006) for further extensions and *NormalShrink* estimates of Kaur et al. (2002).

Image data is often corrupted by noise during its acquisition or transmission. Image denoising is traditionally performed in the spatial or frequency domain by filtering, and is used to remove the additive noise, while retaining, as much as possible, the vital features of the given image. Since Donoho and Johnstone, there has been much investigation on how to optimally define threshold levels and their type (i.e. hard or soft thresholds) (see Bruce & Gao, 1995). Resultant algorithms generally perform a global thresholding of the wavelet coefficients, which involves retaining only large coefficients and putting the rest to zero. As such, they are not spatially adaptive andconsequently their performance in image analysisis not adequately effective (Ellinas et al., 2005).

There is a large body of research that demonstrates that wavelets provide an appropriate basis for separating the noisy signal from the image signal. As such, there has been a considerable amount of work done on filtering and wavelet coefficients thresholding.These wavelet-based methods depend mainly on thresholding the DWT coefficients, which are assumed to have been affected by Additive White Gaussian Noise (AWGN). Indeed a broad class of image processing algorithms is based on the DWT, with the advantage that the transform coefficients within the subbands can be modelled locally as independent identically distributed (i.i.d) random variables with GGD (Mallat, 1989). As such, the denoised coefficients can be calculated by an MMSE (Minimum Mean Square Error) estimator, written as the noise coefficients and the variances of signal and noise. The signal variance is estimated by a Maximum Likelihood (ML) estimator or a Maximum A Posteriori (MAP)) estimator in small regions for every subband (where variance is assumed nearly constant); whereas the noise variance is estimated from the first level diagonal details. The denoised coefficients are therefore estimated statistically in small regions for every subband as an alternative to use a global threshold (Mihcak et al., 1999). In Chang et al. (2000) a comparable model, which is spatially adaptive, and based onwavelet image coefficients was used to execute image denoising via wavelet thresholding.

Later Ellinas et al. (2005) employed the spatially adaptive model described in Mihcak et al. (1999) and performed MMSE coefficient estimation, instead of coefficient thresholding (as used by Chang et al. (2000b)). The way that the underlying variance field is estimated in Ellinas et al. (2005) differs from that discussed by Mihcak et al. (1999). More specifically, in Ellinas et al. (2005) this estimation was performed in a variable block size framework but in Mihcak et al. (1999), a fixed block size framework was employed.

In regard to various Bayesian block thresholding applications: we note that Antoniadis et al. (2002) proposed a framework in which sparseness can be expressed easily and naturally by a

Bayesian model for the wavelet coefficients of the underlying signal. This Bayesian formulation was based on the empirical observation that the wavelet coefficients can be summarised adequately by exponential power prior distributions. Thereby close connections between wavelet thresholding techniques and MAP estimation were shown for two classes of noise distributions including heavy-tailed noises; evidence that a large assortment of thresholding rules can be derived from such MAP criteria. Abramovich et al. (2002) proposed an empirical Bayes approach to incorporating information about neighbouring empirical wavelet coefficients into function estimation which lead to block wavelet shrinkage and block wavelet thresholding estimators. These estimators compared favourably with estimators obtained via several then current non-Bayesian block wavelet thresholding methods (see the references given above).

1.4 Detecting Change Points by Wavelets

Change point detection is the identification of abrupt changes in the generative parameters of sequential data. Detecting change points in time series data is an important data analytic task with application to various scientific domains, for example in bioinformatics and proteomics; in detecting gene sequence changes associated with cancer and in the analysis of Human immunodeficiency virus (HIV) marker responses (Gosh & Vaida, 2007; Khattree & Naik, 2008), a method of wavelet change-point prediction to transmembrane proteins by Lio and Vannucci (2000), Aroian & Levene (1950) developed into a fundamental problem in the area of statistical control theory *per se*. Change point detection via singular spectrum analysis has also been proposed recently application (Hudson & Keatley, 2010b).

The occurrence of change points often reveal important information about the object under study, hence the great interest in detecting and locating change-points (say, in a given function). Generally the main goal of change point problems is the estimation of the number, locations and size of possible change points (Müeller (1998)). Müeller (1998) reviewed some basic results and discussed some pertinent issues regarding the modelling of discontinuous phenomena, which invariably embrace change-point detection. As the unknown function under investigation may contain one or more sudden localized changes, a spatially adaptive method is needed. Since a jump or sharp cusp in the underlying function results in several wavelet coefficients (which are adjacent each other), the theory of wavelets permits decomposition of functions into localized oscillating components. This gives the motivation to turn to the study of localized changes through wavelets as an ideal tool.

As an overall summary of the development of change point methods we note the following. The methods of change point estimation have been considered in Gijbels et al. (1999), Huh and Park (2004), and by Korostelëv & Tsybakov (1993). Detecting change-points in non-parametric regression has received much attention of late; see the works of Antoniadis and Gijbels (2002), Gijbels and Goderniaux (2004, 2005), Raimondo and Tajvidi (2004), and Huh and Park (2004). Research on change-points in non-parametric density estimation includes the work of Neumann (1997) and Huh (2002). Other recent developments on change-point methods include Park and Kim (2004) and Reiss (2004). Other work on change-points in non-parametric regression are the developments described by Müeller (1992), Korostelev and Tsybakov (1993), Cline et al. (1995), Gijbels et al. (1999) and Huh and Carriere (2002). The estimation of the location of so-called jump points (referred to as kinks) in the first derivative of a regression function μ, was very recently studied by Wishart and Kulik (2010) in two random design models with differing long-range dependent (LRD) structures. Their method involves the zero-crossing technique and uses high-order kernels. Knowledge of change points will allow us to identify change in trends in the underlying regression function of a non-parametric model. This thus could explain the change in qualitative or quantitative behaviour of an underlying process. See also the exposition of Wishart (2009). The zero-crossing change point technique has been applied earlier by Cheng and Raimondo (2008) to estimate a kink, instead of a jump point.

We now review, in more detail, paper by paper, the development of change point methods via wavelets, particularly using MRA methods (Mallat, 1989) and the DWT (Nason and Silverman, 1994). Abramovich et al. (2000) presents a good earlier review of wavelets methods to change point analysis, especially the well-established applications of wavelets in statistics including their use in nonparametric regression, density estimation, inverse problems, change point problems and also to some specialized aspects of time series analysis. See also the review of Chen and Gupta (2001) who provide a survey of the change point detection and estimation; the origin of the change point problem; methods for determination of the presence of a statistically significant change point in a sequence of chronologically ordered data via the likelihood-ratio procedure, the informational approach and the Bayes solution.

Wavelet bases offer a measure of localisation in space as well as in frequency that allows for decomposition of a signal into compactly supported oscillating components (Raimondo & Tajvidi, 2004). The coefficients associated with each of these components are called wavelet coefficients or wavelet transforms. A noteworthy property of wavelet coefficients is that they reflect the local regularity of the original function. Recall that they are large where the function is irregular, and small where the function is smooth. This property is very helpful in the detection of discontinuities or of sharp changes in a signal with noise (Raimondo and Tajvidi 2004). See also

Raimondo (1998) and Raimondo (2002). In regard to the wavelet detection of jump and sharp cusps Wang (1995) proposed a test statistic based on the optimisation of the absolute value of the wavelet coefficients. Later Odgen and Parzen (1996a) developed an approach based on the cumulative sum of squared wavelet coefficients. These techniques all detect a "jump" or "cusp" in an assumed differentiable function, which is observed with noise.

Furthermore, Odgen and Parzen (1996b) developed and adapted wavelet thresholding and non-parametric regression to the change point problem. Odgen and Parzen (1996b) provided a data dependent technique for selecting a threshold by which to shrink the empirical wavelet coefficients. Their technique, as based on standard statistical tests of hypotheses, was shown to give first-class results both in the scenario where the underlying function is constant, and when it exhibits multiple abrupt changes. By adjusting the level of the tests of significance, they showed that the smoothness of the resulting estimator can be controlled, allowing preference to be given to good expected mean square error (MSE) performance.

Later Raimondo (1998) defined the sharp change point problem as an extension of the above-mentioned earlier problems in change point analysis related to nonparametric regression. Raimondo (1998) pointed out that estimation of jump points in smooth curves are specific cases which may be included within his paradigm. Raimondo (1998) also provided sensible examples of thresholding empirical wavelet coefficients so as to accurately estimate the position of change points that are sharp. Around the same time as the work of Raimondo (1998) the estimation of change-points (and boundary curves) to testing for the presence of change-points and determination of the number of change-points was shown by Müeller (1998) to fall within the problem of the non-parametric analysis of functions which combine both discontinuous and smooth features. Mueller (1998) demonstrated that various popular change-point detection methods, to that date, and still, namely wavelets, semi-parametric modelling, one-sided kernel differences, and local polynomial fitting, are just different versions of the same fundamental concept, namely that of one-sided kernel estimators. See also the kernel-type diagnostics of Gijbels et al. (1999) for the estimation of jump points in smooth curves, which utilised a non-parametric regression approach and specifically a two-step method based on kernels and then on a local least-squares approach.

The approach of a segmented MRA to change point detection and the estimation of the location of discontinuities in one-dimensional piecewise smooth regression functions (observed in white Gaussian noise over an interval) was utilised by Antoniadis & Gijbels (2002). Their method employs a wavelet analysis of the observed signal and belongs to the class of so-called indirect methods, where the change points are located prior to curve fitting. Subsequently one can then

employ their favourite function estimation technique on each segment (according to the located change points) to model the whole curve. Antoniadis and Gijbels (2002) showed that, provided discontinuities can be detected and located with sufficient accuracy, detection followed by wavelet smoothing indeed has optimal rates of convergence.

Raimondo and Tajvidi (2004) later suggested a method based on checking the *absolute value* of wavelet coefficients which then permits the detection of discontinuities in acurve, even when additive noise is present. Specifically Raimondo and Tajvidi (2004) combined wavelet methods with extreme value theory to test for the occurrence of an arbitrary number of discontinuities in an unknown function (observed with noise). Raimondo and Tajvidi's (2004) approach was based on a Peaks Over Threshold modelling approach of noisy wavelet transforms, which estimates an extreme value index in the tail of the noise, for a critical region derived assuming a Generalised Pareto Distribution (Hosking & Wallis, 1987; Pickands, 1975) approximation to normalised sums. Asymptotic results illustrate that their method was powerful for a wide range of medium size wavelet frequencies.

Park and Kim (2004) proposed a sharp change point estimator based on the differences between right and left boundary wavelet smoothers, constructed by applying a two-step procedure to the observed data. This estimator was shown to enjoy a minimax (Raimondo 1998) convergence rate. Park and Kim (2004) then estimated the regression function with boundary wavelets in the left and to the right regions of the estimated jump point separately. Both mean integrated squared error and mean squared errors of the estimated function were derived. These same authors later developed wavelet estimation of a regression function with a sharp change point in a random design (see also Wishart and Kulik, 2010). Specifically Park and Kim (2006) furthered the block wavelet thresholding approach of Park and Kim (2004) as follows. In a random design nonparametric regression model, their method was based on a design transformation and binning to convert a random design into an equally spaced design whose number of points is a power of 2. Using the continuous wavelet transform (CWT) of the data, they constructed a sharp change point estimator and obtained its rate of convergence.

In summary: the general scheme used to detect a function's abrupt changes through a wavelet approach is based on the link between the function's local regularity properties at a certain point and the rate of decay of the wavelet coefficients near this given point across increasing resolution levels (see, for example, Mallat and Hwang (1992), and Ogden (1997)). Local singularities are identified by abnormal behaviour in the wavelet coefficients at high-resolution levels at the change point. Such concepts are discussed in greater detail in Wang (1995, 1998), Raimondo (1998, 2004) and Park and Kim (2006). In these studies, both fixed and the more complex random

design problem are discussed. The computational results from Wang (1995, 1998), Raimondo (1998), Cai (2002), Delouille et al. (2001), and Park and Kim (2006) all demonstrate superior performance of wavelet estimates above alternative approaches. This is true in both the detection of change points and in the estimation of the regression functions.

In Chapter 6 we develop a novel change point approach. This uses MODWT coefficients to link to a shifting DWT (SDWR), which were used successfully to detect a change point in a time series. Our SDWT approach is a combination of DWT and MODWT technique for the change point detection problem.

1.5 Review and Motivation of Agitation-Sedation (A-S) Modelling

Critically ill patients are admitted to the intensive care unit (ICU) for life support and specialised treatment. Patients' pain and discomfort can lead to anxiety that adds significantly to the patient's agitation and possibly to increased heart rate and blood pressure. This agitation reduces the ability of the patient to recover (Chase et al. 2004). The length of stay in the hospital is thus often increased as are other potential risks as well (Kress, Pohlman, & Hall, 2002). Recent studies have highlighted the cost and health care benefits of drug delivery protocols based upon sedation assessments scales (Brattebo et al., 2002; Kress et al., 2002). The use of quantitative modelling to enhance understanding of the A-S system and the provision of an A-S simulation platform are key tools in this area of patirnt critical care. Several recent quantitative models of the agitation-sedation system have been developed and all attempt to model the effect of sedative and analgesic drugs on patuent agitation in the ICU (Lee et al., 2005; Rudge et al., 2006a, 2006b, 2003; Shaw et al., 2003).

We now closely follow the development of the mathematics in Chase et al. (2004). The first application in this thesis uses two models from the above listed literature on agitation-sedation modelling to illustrate the wavelet methods proposed. These are described in detail in Chapters 2 and 3 (see Figure 2.3 and Table 3.7). The first model used (see Chapter 2) is based on the simulation equations of Chase et al. (2004) which were shown to capture the fundamental dynamics of the agitation-sedation system, but to lack physiological resemblance and more advanced dynamics. Table A.1 in the appendix gives an overview of recent ICU agitation studies and also provides a brief overview of the equations used for simulations of A-S status and the methods derived in this thesis and by other authors with the aim of establishing the validity of the models in reflecting a patient's true agitation-sedation status. Time series profiles of A-S status per patient were very long in length (see Table A.3 in Appendix A). Chase et al.'s (2004) model

builds upon a well-known general two-compartment pharmacokinetic (PK) model (Wood & Wood, 1990), adding patient agitation as a third state variable (see the Table A.1 in the appendix).

The second model (Rudge et al., 2006a) used in Chapter3 adds more complex dynamics to the model, capturing synergism, saturation, and a non-linear concentration-effect dynamics. This model captures both the fundamental dynamics of the agitation system captured in the initial mode, while at the same time being more representative of the physiology (Rudge et al., 2006a, b). Rudge's 2006 model utilises separate pharmacokinetic models for two drugs, midazolam (a sedative) and morphine (a strong analgesic with mild sedative effects), together with a drug response surface to represent the combined sedative effect of these drugs. The model also provides a close representation of the actual physiological system and includes complex dynamics, such as delayed distribution, drug synergism and effect saturation (Rudge et al., 2006a, 2006b). The model defined and quotes directly from Rudge et al, (2006b) in three main portions which was shown in the Table A.1. Details of the development and validation of the agitation-sedation system model can be found in Rudge et al. (2006a,2006b), Chase et al.(2004) and Shaw et al (2006).

Numerical and graphical approaches using wavelet methods are then used to provide statistical measures of wavelet time coverage index (WTCI) (see Chapters 2 and 3 of this thesis) and to verify the model's ability to capture the fundamental dynamics of the A-S system.

1.6 Review and motivation of the neuroinformatics and psychometric modelling application: Personality and Brain function

Personality has been of interest in the psychometric literature for at least 3 decades (see Table A.2 in the appendix, which is sourced from Hermes, 2007) for an overview of research into personality). A wide variety of aspects have been studied. A variety of personality models are used these days, each with their own advantages and disadvantages. A recently developed model that offers the advantages of a biological basis is the temperament and character inventory (TCI) model of Cloninger (1994) (see also Cloninger 2002, 2003, 2008). This model introduces seven basic traits that measure unique personality axes. These traits are novelty seeking (NS), harm avoidance (HA), reward dependence (RD), persistence (P), self directedness (SD), cooperativeness (C) and self transcendence (ST). These seven traits are categorised into either temperament or character descriptors. The temperament descriptors (novelty seeking, harm avoidance, reward dependence, persistence) are moderately heritable and stable (Cloninger, 1994).

Character (self directedness, cooperativeness and self transcendence) on the other hand is described by Cloninger et al. (Cloninger, 1994, 2003) as referring to the " self concepts and individual differences in goals and values that influence voluntary choices, intentions and the meaning of what is experienced in life". Thus unlike temperament, character changes with age and is influenced by social experiences.

Increased understanding of the neurobiology and the neuroanatomy of personality traits, has been much progressed by recent advances in brain imaging technology. Importantly this technology is also adding to our understanding of fundamentals such as: learning (Haier, 2001; Haier et al., 1992), memory (Alkire et al.,1996, 1998), intelligence (Haier et al., 1988; Haier, White, & Alkire, 2003), and consciousness (Alkire & Haier, 2001; Alkire, Haier, & Fallon, 2000). Understanding the relationship between activity in specific areas of the brain and certain mental processes, relies heavily on measuring brain function.This is commonly achieved through functional neuroimaging, which is a popular research instrument in the associated areas of cognitive neuroscience and neuropsychology (Friston et al., 2007; Frackowiak et al., 1997; Frackowiak et al., 2001; Kolb, 1990; Farmers & Goldberg, 2008; Bullmore et al., 2001, 2003; Canli, 2001; Desco et al., 2001).

Various methods exist for deriving a map of the functional activity in the brain. One can, for example, measure localised changes in cerebral blood flow, which are presumed to be associated with neural activity. Termed 'activations', these are caused by regions of the brain being activated by a specific task being performed by the subject (person). Activations and knowledge of their cerebral locations help our understanding of the neural computations which contribute to a given behaviour. For example, when a subject performs tasks, which involve specifically visual stimulation, it is usual to observe extensive activation of the occipital lobe, which is a section of the brain which receives signals from the retina. This region of the brain is then assumed to play a crucial role in visual perception.

PET, fMRI and SPECT are popular and recently developed non-invasive procedures that measure regional cerebral blood flow (rCBF) changes. Researchers in personality have however, limited access to imaging technology, given the high budget needed for such collaborative research (often across other specialties). Scanning of large samples with tasks judged to be appropriate to issues in personality research would be required, along with image analytic methods that could address the dual problems of anatomical localisation and statistical inference. Despite the many difficulties, various brain imaging techniques have successfully identified the functional neuroanatomy of personality traits. As computing technology improved and voxel-by-voxel analysis became achievable, however, this problem grew even more difficult and required alternative statistical approaches (Zuckerman & Stelmack, 2004).

Standard image analysis programs like Statistical Parametric Mapping (Friston et al., 1995b; Friston et al., 2003, 2007) are now regularly utilsed in modern studies into personality. The technique, as implemented in SPM5 which uses the general linear model (Kiebel & Holmes, 2003; Frackowiak et al., 1997; Friston et al., 1995b) at each and every voxel simultaneously to construct a map characterising the null hypothesis at each voxel. This map is assessed for significant voxels or regions. (http://www.fil.ion.ucl.ac.uk/spm/software/spm5).

Chapter 5 presents an independent components analysis using FastICA (Hyvarinen, 1998) combined with statistical parametric mapping (SPM) (Friston, Ashburner, & Heather, 2003) as an approach by which to model the brain function and personality data from Turner et al (2003) (see Figure 1.1 and Table 5.2). ICA is a statistical and computational technique for revealing hidden factors that underlie sets of random variables, measurements, or signals (Hyvärinen et al., 2001b) and is used in Chapter 5 primarily to reduce noise in the 3D brain images. See Figure 1.1 which illuatrates the links between chapters of this thesis. The theory behind FastICA is published in Hyvärinen, (2001a) and Hyvärinen et al. (2001b). FastICA, FA and PCA are part of the broader hierarchy of ICA models see Page 7 of Roberts and Everson, 2001). In chapter 4, SPECT images were used for denoising and also in the TCI model and brain function application discussed in Chapter 5. The differences in regional blood flow are investigated in relationship to differences in personality types, using functional brain imaging and ICA. Personality is measured using the TCI (Cloninger, 1994, 2002, 2003, 2008) and brain function is measured using SPECT (Prohovnik, 1993).

1.7 Overview of the thesis

This section gives an overview the main aims of the thesis and describes the schema in Figure 1.2 of the methods and applications of the thesis. The arrows in Figure 1.2 show significant interconnections between the wavelet methodologies as applied to Chapters 2-6 of the thesis. Each chapter is also summarised below specifically in regards to the methodology proposed and the medical application used, namely, one of the following; [1] the agitation sedation time series data for 37 ICU patients (also briefly described in Section 1.2 of this Chapter) and [2] the neuroinformatic data from Turner et al (2003) in a study which aims to link personality with brain function (see also Section 1.3 of this Chapter for a brief description and motivation of the brain image and psychometric data). This research presented in this thesis has five primary aims:

1. To develop wavelet modelling of bivariate time series based on the concepts of wavelet correlation and wavelet cross correlation. These methods are applied to an ICU based A-S

system (data based on the study of Chase et al. (2004). The objective of aim is to increase understanding and possibly progress steps to create platforms enabling the development of agitation feedback protocols for patients in ICU.

2. To develop a nonparametric approach based on wavelet smooth regression and density estimation for assessing the validity of the deterministic dynamic A-S models (see 1. above) against empirical data (see Figure 1.2). A secondary aim is to obtain visual graphical assessment tools as well as numerical metrics of compatibility between the simulated (modelled) and the observed A-S time series data.

3. To compare several wavelets methods by which to denoise images, specifically SPECT brain images (used in the psychometric and brain study of Turner et al. (2003) and analysed in Chapter 5 of this thesis).

4. To apply ICA and SPM to model SPECT) brain images with respect to psychometric constructs of personality.

5. To develop a novel wavelet method of detecting the location of a change point in a time series, using wavelet transforms and shifts.

The statistical analyses using wavelets and as applied to two medical data sets are introduced and discussed in the following chapters of this thesis. Figure 1.2 gives an overview schema of the methods and applications of the thesis. The arrows show significant interconnections between the wavelet methodologies developed.

Chapter 2 presents a wavelet analysis to A-S modelling. Section 2.2 gives a brief literature review of the wavelets. Section 2.3 presents the application of agitation-sedation wavelet modelling using wavelet-based diagnostics. Specifically in Chapter 2 a suite of wavelet analytic techniques are used which are based on the discrete wavelet transform multiresolution analysis (DWT-MRA) and maximal overlap discrete wavelet multiresolution analysis (MODWT-MRA) by which to assess whether an ICU patient's simulated agitation- sedation status over time relects their true dynamic (A-S) profile. Simualted profiles are based on differential equations and non-parametric regression methods as derived by Chase et al. (2004) and Rudge et al. (2006a) (see Table A.1 in the appendix). This chapter builds on initial work by Hudson et al. (2004), which was a preliminary study to assess wavelet signatures for modelling ICU agitation-sedation profiles, and to, as in Chapter 2, evaluate "closeness" or "discrimination" with respect to wavelet scales. Another earlier application was by Kang et al. (2005) on an australian phenological climatic time series study where it was established that wavelets add credibility to the use of phenological records to detect climate change, recently reported by Hudson et al. (2010d) and due for publication by Hudson et al. (2010a) (see also Hudson et al., 2003; Keatley et al., 2005). To the best of our knowledge, the work in Chapter 2 represents the first application of wavelet

cross-correlation based wavelets analysis of agitation-sedation dynamics (see Figure 1.2 and Table A.1 in the appendix).

In Chapter 3 a nonparametric approach for assessing the validity of the deterministic dynamic A-S models (see the ICU data discussed in Chapter 2) against empirical data is developed (see Figure 1.2 and Table A.1 in the appendix). The approach in Chapter 3 is based on wavelet smooth regression and its density estimation, yielding visual graphical assessment tools as well as numerical metrics of compatibility between the simulated (modeled) and the observed A-S time series data. Specifically a Bayesian approach is suggested in Chapter 3 by which to assess a parametric A-S model – this by constructing a wavelet probability band (WPB) for the proposed model and then checking whether the *nonparametric* regression curve lies within the band. Our contribution in Chapter 3 is to construct a wavelet probability band for the nonparametric wavelet regression curve and check whether the proposed model lies within the band. The wavelet probability band (WPB) also gives a useful tool to measure compatibility of simulated and recorded time series profiles. Moreover, the density profile is used to define and compute two numerical measures, namely the average normalized wavelet density (ANWD) and relative average normalized wavelet density (RANWD); measures of agreement between the recorded infusion rate and simulated infusion rate. Section 3.1 describes the basic smoothing techniques based on wavelets and methods by which the density is estimated using wavelet shrinkage methods, as described in Section 3.2. In Section 3.2, wavelet shrinkage based on Bayesian method is applied to the ICU data. Specifically in Section 3.2.1.1, we use the minimax estimator to obtain a patient specific WTCI. All values of the WTCI are obtained using Bayesian wavelet thresholding (Nason, 2008). As far as the author is aware this is the first study to develop a nonparametric wavelets approach for assessing the validity of deterministic dynamic A-S models against empirical data, based on wavelet smooth regression and wavelet density estimation. Equally important are the resulatant visual graphical assessment tools, in addition to the wavelet based numerical metrics for compatibility between the modelled and observed A-S data (see Table A.1 in the appendix).

Chapter 4 considers wavelets based brain image denoising techniques. Section 4.2 presents the background for 2-D image theory and compares four different denoising methods in Section 4.3. In Chapter 4 two different wavelet transform schemes are designed, namely the DWT and MODWT, with four different wavelet bases. Their performance for image denoising is compared in Section 4.3.2. The aim of our designs is to study the suitability of different wavelet bases and also study the effect of different window sizes. In Section 4.3.3 we compare wavelet denoising versus ICA (see Figure 1.1).

In the application of Chapter 4, denoising of images is restricted to 2D. However, SPECT brain image is 3D data, and denoising is generally performed in 3D space to take advantage of the resultant better separation of noise and signal in higher dimensions and the availability of volumetric features available in true 3D datasets.

In Chapter 5, we use SPECT 3D brain image data to remove noise by using independent component analysis (see Figure 1.2 which illustrates the links between the chapters of the thesis). Specifically Chapter 5 presents an independent components analysis using FastICA (1998b), combined with statistical parametric mapping (SPM) (Friston, Ashburner, & Heather, 2003) as an approach by which to model brain function and personality (see Figure 1.2 and Table A.2 in the appendix). Details on the procedural steps and mathematics of ICA and on FastICA are given in Chapter 5. ICA is used in this thesis principally to denoise the brain image data, prior to modelling brain activations or deactivations, with respect to the psychometric constructs of personality and character (Cloninger, 1994, 2003). Section 5.3 presents an overview of the psychometric and brain function data and constructs, and Section 5.4 describes the SPM (Frackowiak et al., 1997, 2003, 2004; Friston et al., 1995a; Friston et al., 1995b; Friston, 2003, Friston, Ashburner, & Heather, 2003) (see also Section 1.3 of this chapter, Chapter 1). ICA theory is presented in Section 5.5 (see also section 1.2.1 of this chapter, Chapter 1) and the main results are given in Section 5.7. This neuroinformatics study in Chapter 5 extends the work of Turner et al. (2003) by using ICA to assess the difference in cerebral blood flow (CBF) between seven temperament and charater, so-called personality traits (Cloninger, 2002, 2008).

In Chapter 6 a novel wavelets based change point approach is developed, which uses MODWT coefficients to link to a SDWT methodology. Our SDWT is thus a combination of the DWT and MODWT for the change point detection problem. Section 6.4 examines some of the existing change point detection methods, namely, Atheoretical Regression Trees (ARTs) and Bai and Perron (Bai & Pearson, 2003; Bai & Perron, 1998) (BP)' s method and wavelet methods (Donoho & Johnson, 1998; Ogden, 1997). These are compared to our results using the SDWT on a simulated data set. As far as the author is aware this is the first time to date that a shifted DWT method has been developed to detect a change point.

Chapter 7 summarises the results presented in the thesis and suggests avenues of future work.

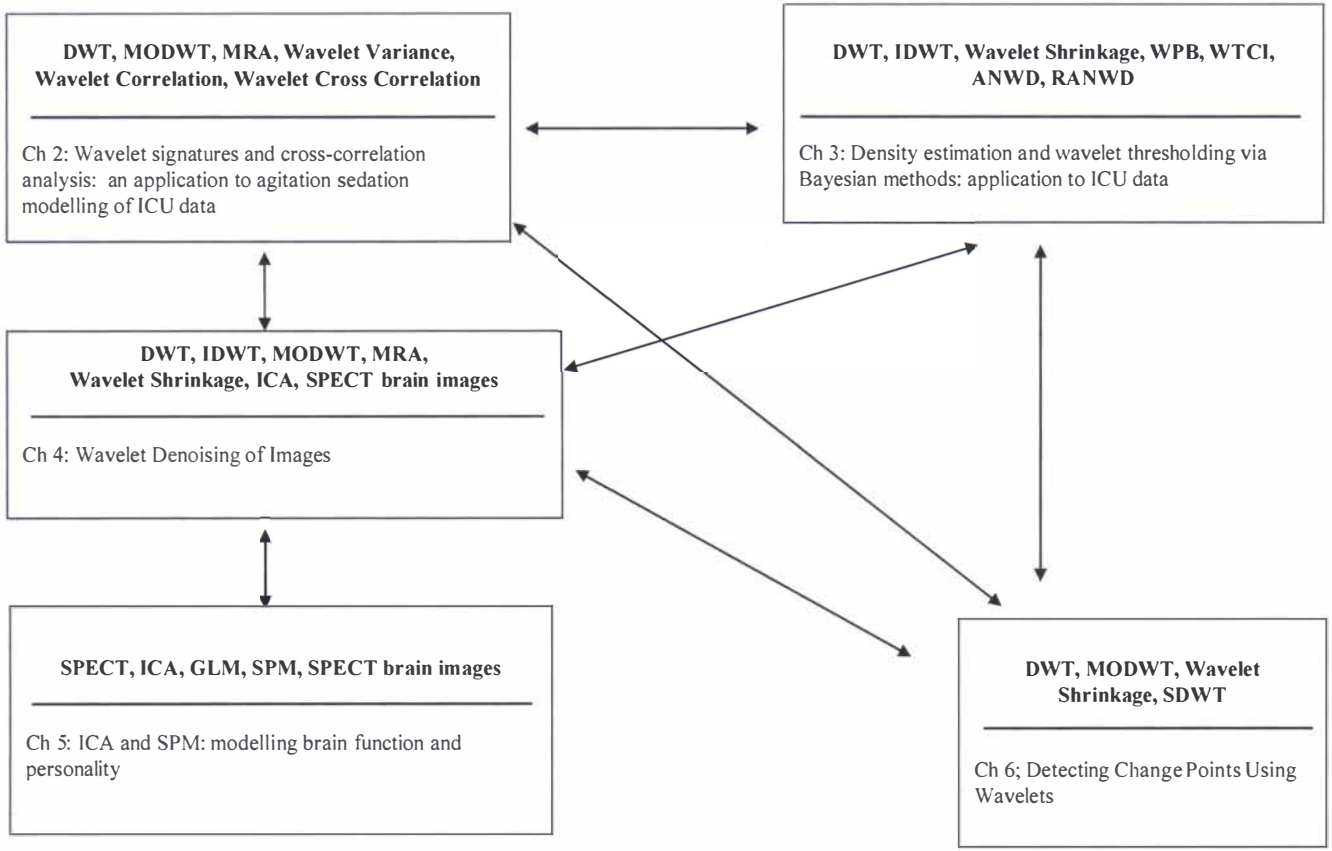


Figure 1.2 Overview schema of the methods and applications of the thesis. The arrows show significant interconnections between the wavelet methodologie

Chapter 2

2 Wavelet Signatures and Diagnostics: for the Assessment of ICU Agitation-Sedation Protocols

2.1 Introduction

A wave is usually defined as an oscillating function that is localized in both time and frequency. A wavelet is a "small wave", which has its energy concentrated in time to give a tool for the analysis of transient, nonstationary, or time-varying phenomena (Goupillaud et al., 1984; Morlet, 1983). Wavelets have the ability to allow simultaneous time and frequency analysis via a flexible mathematical foundation. Wavelet is well suited to the analysis of transient signals in particular. The localizing property of wavelets allows a wavelet expansion of a transient component on an orthogonal basis to be modelled using a small number of wavelet coefficients using a low pass filter (Barber et al., 2002). This turns out to be applied in a wide range of fields, such as signal processing, data compression, and image analysis (Mallat, 1998; Meyer, 2003; Kumar, 1993, 1994; Donoho, 1995; Chang et al., 2000a, 2000b).

Typically agitation-sedation cycling in critically ill patients involves oscillations between states of agitation and over-sedation, which is detrimental to patient health, and increases hospital length of stay (Rudge et al. 2006a; 2006b; Chase et al 2004; Rudge et al 2005). The goal of the research by Chase et al. (2004), was to develop a physiologically representative model that captures the fundamental dynamics of the agitation-sedation system. The resulting model can serve as a platform to develop and test semi-automated sedation management controllers that offer the potential of improved agitation management and reduce length of stay in the intensive care unit (ICU). Chase et al. (2004) presented a minimal differential equation model to predict or simulate the patients' agitation-sedation status over time, which captured the agitation-sedation dynamics of 37 ICU patients.

In this thesis closeness between a patient's simulated and mean recorded actual series was quantified in two ways. Firstly by the calculation of the wavelet coverage percentage of time a patient's simulated profile fell within a 90% wavelet probability band (Barber et al., 2002) of their mean infusion rate (Chase et al., 2004). Wavelet regression using shrinkage (threshold) was used to obtain the ICU patients' mean infusion rate. Secondly, by the calculation, via block

bootstrapping of time series, a patient-specific wavelet time coverage index (WTCI) between the patient's simulated and estimated mean infusion rate was calculated. 100% tracking was indicative of simulated and mean profiles being "close" or synchronous.

Recently a more refined model, which utilises an Epanechnikov kernel, captures the fundamental agitation-sedation (A-S) dynamics was formulated by Rudge et al. (2006a; 2006b). The work in Chapter 2 investigates the use of wavelet signatures and statistics on Rudge et al's (2006a; 2006b) simulated profiles, to test for commonality across patients, in terms of wavelet correlations. A secondary aim of this study is to test the feasibility of wavelet statistics to help distinguish between patients whose simulated profiles were "close" to their mean profile versus those for whom was not the case (i.e. their simulated profiles are not "close" to their actual recorded profiles).

This chapter builds on initial work by Kang et al. (2005), which was a preliminary study to assess wavelet signatures for modelling ICU agitation-sedation profiles, so as to, as in this chapter, evaluate "closeness" or "discrimination" of simulated versus actual A-S profiles with respect to wavelet scales. Another earlier application of some of our methods was the study by Kang et al. (2004) on an Australian phenological flowering climatic time series, where it was established that wavelets add credibility to the use of phenological records to detect climate change. This study was recently expanded and reported by Hudson (2010c), Hudson and Keatley (2010a) (see also Hudson et al., 2004; Hudson et al., 2005; 2010c, 2010d).

2.2 Brief Literature Review

This chapter gives a brief introduction of the basic ideas concerning wavelets. The wavelet decomposition of functions is related to analogous Fourier decomposition methods (Ogden, 1997; Abramovich & Benjamini, 1995). The wavelet representation is presented first in terms of its simplest paradigm, the Haar wavelet (Haar, 1910). The Haar wavelet is used in this chapter to describe the concepts of multiresolution analysis (MRA), and these ideas are generalized to other types of wavelets. This chapter provides some of the background needed for the rest of the research and application in this thesis except Chapter 5.

For more details about wavelets see, for example, Daubeches (1992), Chui (1992), Donoho & Johnstone (1994), Ogden (1997), Vidakovic (1999), Percival & Walden (2000), and Gencay et al. (2001). There is a great amount of statistical literature on wavelets. Most of the work in developing the mathematical foundations of wavelets has been completed (see Chapter 1 for relevant literature). It remains for us to study their applications in various areas, our work covers

wavelets based diagnostics for the assessment of ICU agitation-sedation protocols in Chapters 2 and 3 of the thesis.

2.2.1 The Discrete Wavelet Transform (DWT)

Wavelets may be formed from the mother wavelet function $\psi(t)$ via dyadic dilation and integer translation by the following,

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \quad j,k \in Z, \tag{2.1}$$

where $\mathbb{Z}$ is the set of all integers and the factor $2^{j/2}$ maintains a constant norm independent of scale j. The entire set of wavelets $\psi_{j,k}(t)$ forms an orthonormal basis (Daubeches, 1992). The wavelet functions $\psi_{j,k}$ are ordered according to their dilation index j and translation index k. Higher j corresponds to lower frequency wavelets, and higher k corresponds to a rightward shift.

The wavelet transform is usually considered to be a continuous wavelet transform (CWT) (Vidakovic, 1999; Percival & Walden, 2000; Gencay et al., 2001) since it is applied to a function $f(\cdot)$ defined over the entire real axis. However, we only have a finite number N of sampled values, as is usually the case for real data applications. This approach leads to the discrete wavelet transform (DWT). To some degree of approximation, we can regard the DWT as being formed by taking slices through a corresponding CWT (McCoy et al., 1995). Any wavelet in $L^2(R)$ then can then be written as a set of expansion functions,

$$f(t) = \sum_{j,k} a_{j,k} 2^{j/2} \psi\left(2^j t - k\right), \tag{2.2}$$

where the two-dimensional set of coefficients $a_{j,k}$ is called the discrete wavelet transform (DWT) of $f(t)$. A more specific form indicating how the $a_{j,k}$'s are calculated, can be written using inner products (Swelden, 1996) as follows,

$$f(t) = \sum_{j,k} \left\langle f(t), \psi_{j,k}(t) \right\rangle \psi_{j,k}(t). \tag{2.3}$$

Let $\boldsymbol{X}_t$ be a dyadic length column vector containing a sequence $X_1, X_2, \cdots, X_N$ of $N=2^J$ observations of a real-valued time series. The length N vector of discrete wavelet coefficients $\boldsymbol{W}$ is obtained via $\boldsymbol{W}= \mathcal{W}\boldsymbol{X}$, where $\mathcal{W}$ is an $N \times N$ orthonormal matrix defining the DWT. The vector of wavelet coefficients may be organised into $J + 1$ vectors,

$$\boldsymbol{W} = \left[W_1, W_2, \cdots, W_J, V_J\right]^T, \tag{2.4}$$

where $\boldsymbol{W}_j$ is a length $N/2^j$ vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and $\boldsymbol{V}_J$ is a length $N/2^J$ vector scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$. Wavelet coefficients thus tell us about variations in adjacent averages (Percival & Walden, 2000).

The structure of the $N \times N$ dimensional matrix $\mathcal{W}$ is visualised through the submatrices $\mathcal{W}_1$, …, $\mathcal{W}_J$ and $\mathcal{V}_J$ (scaling coefficient matrix) via

$$\mathcal{W} = [\mathcal{W}_1, \mathcal{W}_2, \ldots, \mathcal{W}_J, \mathcal{V}_J]^T \tag{2.5}$$

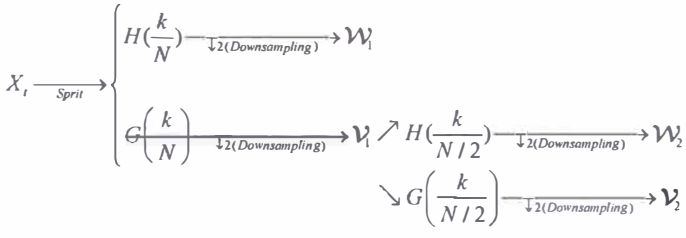


Figure 2.1 Flow diagram illustrating the decomposition of X_t into first and second level wavelet coefficients $\mathcal{W}_{1,t}$ and $\mathcal{W}_{2,t}$ and their scaling coefficients $\mathcal{V}_{1,t}$ and $\mathcal{V}_{2,t}$ (k=0,…,N-1).

Let us now consider implementation of the DWT by using a pyramid algorithm (PA) (Mallat, 1989). Let h = (h_0, . . . , h_{L-1}) be the vector of wavelet (high -pass) filter coefficients and g = (g_0, . . . , g_{L-1}) be the vector of scaling (low-pass) filter coefficients (Daubechies, 1992). Graphical representation of the DWT as applied to a dyadic length vector X_t is given by Figure 2.1 and Figure 2.2. These depict the analysis of X_t into $\mathcal{W}_1$, $\mathcal{W}_2$ and $\mathcal{V}_2$ using the pyramid algorithm (PA). The synthesis of X_t from $\mathcal{W}_1$, $\mathcal{W}_2$ and $\mathcal{V}_2$ use the inverse of the PA. The Inverse DWT (IDWT) is

achieved through upsampling the final level of wavelet and scaling coefficients, convolving them with their respective filters and adding up the two filtered vectors. Figure 2.2 gives a flow diagram for the reconstruction of X_t from the second level wavelet and scaling coefficient vectors.

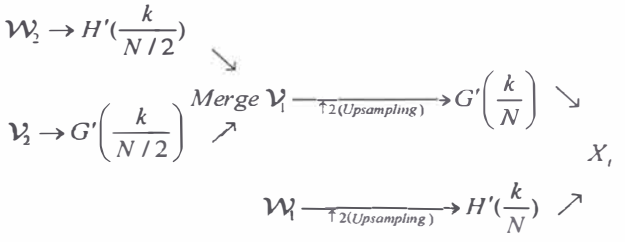


Figure 2.2 Flow diagram illustrating the reconstruction of X_t from first and second level wavelet coefficients $\mathcal{W}_{1,t}$ and $\mathcal{W}_{2,t}$ and their scaling coefficients $\mathcal{V}_{1,t}$ and $\mathcal{V}_{2,t}$ (k=0,...,N-1).

Given a dyadic length time series, it was not necessary to implement the DWT down to level $J = \log_2(N)$. A partial DWT (PDWT) may be performed instead that terminates at a level $J_p < J$. The resulting vector of wavelet coefficients will contain $N - N/2^{J_p}$ wavelet coefficients and $N/2^{J_p}$ scaling coefficients (Percival & Walden, 2000; Gencay et al., 2001). PDWT's are more commonly used in practice than the full DWT because of the flexibility they offer in specifying a scale beyond which a wavelet analysis into individual large scales is no longer of real interest. A PDWT of level J_0 allows us to relax the restriction that $N = 2^J$ for some J and replace it with the condition that N can be an integer multiple of 2^{J_0} (Percival & Walden, 2000; Gencay et al., 2001).

2.2.2 The Maximal Overlap Discrete Wavelet (MODWT)

The DWT is an alternative to the Fourier transform (FT) for time series analysis. The DWT provides wavelet coefficients that are local in both time and frequency. In this section the maximal overlap DWT (MODWT) which is a modified version of the discrete wavelet transform is discussed. Like the DWT, the MODWT is defined in terms of a computationally efficient pyramid algorithm. The term MODWT comes from the relationship of the MODWT with estimators of the Allan variance (Allan, 1966). The MODWT gives up orthogonality in order to gain features the DWT does not possess. A consequence of this is that the wavelet and scaling coefficients must be rescaled in order to retain the variance preserving property of the DWT (Percival & Guttorp, 1994). The properties in Table 2.1 are important in distinguishing the MODWT from the DWT (Percival & Mofjeld, 1997; Percival & Walden, 2000; Gencay et al., 2001).

The decomposition and reconstruction procedure and inverting of the MODWT is similar to the DWT. A key feature to an MRA using the MODWT is that the wavelet details and smooth are associated with zero-phase filters. Thus, interesting features in the wavelet details and smooth may be perfectly aligned with the original time series. This attribute is not available through the DWT since it subsamples the output of its filtering operations (Percival & Walden, 2000; Gencay et al., 2001).

Table 2.1 Properties of the DWT and MODWT

Property	DWT	MODWT
Data	$N=2^J$	Any sample size N
Detail and Smooth Coefficients of MRA	Downsampling	Associated with zero phase filters
Circularly shifting	Do not hold	Holds and Invariant
Wavelet Variance	Less efficient	Asymptotically Efficient

2.2.3 Wavelet-Based Estimators of Covariance and Correlation

The length N vector of discrete wavelet coefficients W is obtained via $\boldsymbol{W} = \mathcal{W}\boldsymbol{X}$ and $\boldsymbol{W} = \left[W_1, W_2, \cdots, W_J, V_J\right]^T$ where $\mathcal{W}$ is an $N \times N$ orthonormal matrix defining the DWT, $\boldsymbol{W}_j$ is a length $N/2^j$ vector of wavelet coefficients associated with changes on a scale of length $\lambda_j = 2^{j-1}$ and $\boldsymbol{V}_J$ is a length $N/2^J$ vector scaling coefficients associated with averages on a scale of length $2^J = 2\lambda_J$ in Section 2.2.1.

Due to the orthonormality, $X = \mathcal{W}^T\boldsymbol{W}$ and the squared wavelet coefficients of $\boldsymbol{W}$ norm, $\|\boldsymbol{W}\|^2$, is the same as $\|X\|^2$ since $\|X\|^2 = X^T X = (\mathcal{W}^T\boldsymbol{W})^T(\mathcal{W}^T\boldsymbol{W}) = \boldsymbol{W}^T\mathcal{W}\mathcal{W}^T\boldsymbol{W} = \|\boldsymbol{W}\|^2$.

Then $\|X\|^2$ is decomposed on a scale-by-scale basis via $\|X\|^2 = \|\boldsymbol{W}\|^2 = \sum_{j=1}^{J}\|\boldsymbol{W}_j\|^2 + \|\boldsymbol{V}_J\|^2$.

Therefore W_n^2 represents variance-preserving transform to the DWT coefficients with index n. The wavelet variance for scale λ_j is defined as the variance of the wavelet coefficients, $W_{j,t}$, then we have the DWT versions of the wavelet variance, that is

$$\tilde{\sigma}_X^2 = E\left(X^2\right) - \left(E\left(X\right)\right)^2 .$$

$$= \frac{1}{N}\left\|X\right\|^2 - \overline{X}^2$$

$$= \frac{1}{N}\left\|W\right\|^2 - \overline{X}^2 = \frac{1}{N}\sum_{j=1}^{J}\left\|W_j\right\|^2$$

where $W_{N-1} / \sqrt{N} = \overline{X}$ (Percival & Walden, 2000; Gencay et al., 2001).

We now define the wavelet correlation and cross-correlation between two time series. The wavelet correlation (WCORR) is the correlation between the scale λ_j wavelet coefficients of bivariate time series. Likewise the wavelet covariance is the covariance between the scale λ_j wavelet coefficients from bivariate time series. We introduce a lag τ, between the two series, to obtain the wavelet cross-covariance (WCCOVA) and wavelet cross-correlation (WCCORR) (Percival & Walden, 2000; Gencay et al., 2001), as described below.

Let X_t and Y_t be two time series of interest. The wavelet cross-covariance of $\left(X_t, Y_t\right)$ for scale $\lambda_j = 2^{j-1}$ and lag τ is defined (Percival, 1995; Percival & Guttorp, 1994; Whitcher et al., 2000; Gencay et al., 2001) as follows,

$$\gamma_{\tau,XY}\left(\lambda_j\right) \equiv Cov\left\{\overline{W}_{j,t}^X, \overline{W}_{j,t+\tau}^Y\right\} , \qquad (2.6)$$

where $\overline{W}_{j,t}^X$ and $\overline{W}_j^Y$ are the scale λ_j MODWT coefficients for X_t and Y_t. When the width of wavelet or scaling filter is equal or greater than two times of the number of differencing operations, the MODWT coefficients have mean zero and therefore the covariance reduces to an expectation of a product ((Percival & Walden, 2000). By setting τ = 0 and X_t to Y_t, $\gamma_{\tau,XY}\left(\lambda_j\right)$ reduces to the wavelet variance for X_t or Y_t denoted by $\sigma_X^2\left(\lambda_j\right)$ or $\sigma_Y^2\left(\lambda_j\right)$. The wavelet correlation of (X_t, Y_t) at scale λ_j=2^{j-1} is defined as follows,

$$\rho_{XY}\left(\lambda_j\right) = \frac{Cov\left\{\overline{W}_{j,t}^X, \overline{W}_{j,t}^Y\right\}}{\sigma_X\left(\lambda_j\right).\sigma_Y\left(\lambda_j\right)} = \frac{\gamma_{XY}\left(\lambda_j\right)}{\sigma_X\left(\lambda_j\right).\sigma_Y\left(\lambda_j\right)} \qquad (2.7)$$

where $\sigma_X^2(\lambda_j) = \text{var}\left\{\overline{W}_{j,t}^X\right\}$ is the wavelet variance with scale λ_j. $\overline{W}_{j,t}^X$ and $\overline{W}_{j,t}^Y$ are the scale λ_j MODWT coefficients for X_t and Y_t, respectively (Percival & Walden, 2000; Gencay et al., 2001). As with the usual correlation coefficient (between two random variables), the range of $\rho_{XY}(\lambda_j)$ is the interval –1 to 1 for all j. The typical cross-correlation statistic is purely a function of the cross-covariance and standard deviations. Thereby the MODWT estimator of the wavelet cross-correlation (WCCORR) of the two processes which are at variance by an integer lag τ, is defined as

$$\rho_{XY,\tau}(\lambda_j) = \frac{Cov\left\{\overline{W}_{j,t}^X, \overline{W}_{j,t}^Y\right\}}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)} = \frac{\gamma_{XY,\tau}(\lambda_j)}{\sigma_X(\lambda_j).\sigma_Y(\lambda_j)}. \tag{2.8}$$

As the usual cross-correlation is used to determine lead or lag relationships between two series, the wavelet cross-correlation should be able to provide a lead or lag relationship between X_t and Y_t on a scale-by-scale basis (Gencay et al., 2001)

Now we introduce confidence intervals (Cls) for the WCORR and WCCORR parameters (Gencay et al., 2001). A nonlinear transformation is required to produce reasonable Cls for the correlation coefficient. We use the Fisher's z-transformation (Kotz et al., 1982), which is defined as follows,

$$h(\rho) = \frac{1}{2}\log\left(\frac{1+\rho}{1-\rho}\right) = \tanh^{-1}(\rho). \tag{2.9}$$

An unbiased estimator of the WCORR based on the MODWT in Equation (2.7) is $\tilde{\rho}$. The given estimated correlation coefficient $\tilde{\rho}$, based on N independent Gaussian observations, has the following limiting distribution,

$$\sqrt{N-3}\left[h(\tilde{\rho}) - h(\rho)\right] \sim N(0,1). \tag{2.10}$$

Table 2.2 Summary of the wavelet mathematics and rationale.

Rationale for wavelets

- Wavelets allow time series data to be decomposed on a scale by scale basis, or discretized, into its underlying subcomponents
- Conventional time frequency domain methods results may be difficult to interpret, whereas the wavelet-correlation displays how the association between two time series change with wavelet scale
- Transformation of the data (orthonormal) allows correlation, cross-correlational analyses of bivariate series to be performed - based on the derived wavelet coefficients
- DWT is often less computer intensive than other transformations (e.g. fast Fourier transform)
- Of the different wavelet methods DWT offers easier analysis than the CWT[++] as most time series are sampled as discrete values
- DWT allows decorrelation of time series

Qualitative Description of the DWT and MODWT

- DWT transforms the original time series $\boldsymbol{X}$ into its DWT coefficients $\boldsymbol{W} = w\boldsymbol{Y}$, where w is a $N \times N$ orthonormal matrix

– the components, $\boldsymbol{W}_j$, of $\boldsymbol{W}$ are associated with coefficients for each scale $\lambda_j = 2^{j-1}$

– Wavelet coefficients $\boldsymbol{W}j$ inform about variations in adjacent averages over λ_j

- There are $N / 2\lambda_j$ wavelet coefficients for each scale $\lambda_j \equiv 2^{j-1}$, $j = 1, 2 \ldots, J_0$ where $\lambda_j = 2^{j-1}$ is a so-called standardized scale, whereas $\lambda_j\Delta t$ is a physical scale, where Δt is the sampling interval
- The MODWT is a non-decimated variation of the DWT.
- MODWT defines the jth level MODWT detail subcomponent of the time series as

  $\overline{D}_j = \overline{W}_j^T \overline{W}_j$ and the jth level MODWT smoothed series, $\overline{S}_j = \overline{V}_j^T \overline{V}_j$,

  which is related to the average (over scale N), normally the trend.

Cross- Correlation and correlation

- The scale λ_j MODWT coefficients may also be used to examine the correlation and cross-correlation of two time series, X_t and Y_t
- The wavelet cross correlation of X_t, Y_t at scale $\lambda_j = 2^{j-1}$ for a time lag τ, is defined as
- $\rho_{xy}(\lambda_j) = \frac{\mathrm{Cov}\{\overline{W}_{x,t}^{\tau}, \overline{W}_{y,t}\}}{\sigma_x(\lambda_j)\,\sigma_y(\lambda_j)} = \frac{\gamma_{xy}(\lambda_j)}{\sigma_x(\lambda_j)\,\sigma_y(\lambda_j)}$, where $\overline{W}_{x,t}^{\tau}$ and $\overline{W}_{y,t}$ are the scale λ_j MODWT coefficients for X_t and Y_t, respectively (Percival, 1995). Correlation is just a case of cross-correlation when $\tau = 0$.

MODWT-MRA

- The orthonormal matrix, $N \times N$, leads to scale-based decompositions
- Given the MODWT coefficients, Y can be constructed as an additive decomposition, known as a multiresolution analysis (MRA).
- Specifically the level J_0 MODWT-based MRA is given by $Y = \sum_{j=1}^{J_0} \overline{d}_j + \overline{s}_{J_0}$. $\overline{d}_j$ are the "detail series", (j= 1, 2, . . ., J_0) for a pre specified J_0, and are part of the MRA of Y that can be attributed to variations on a scale of λ_j .

[+] CWT denotes the continuous wavelet transform

The factor $\sqrt{N-3}$ provides a better approximation of the standard normal distribution (Percival & Walden, 2000; Gencay et al., 2001). Table 2.2 gives a summary of the wavelet mathematics and rationale used in this chapter. When we substitute $h\left[\tilde{\rho}_{XY}\left(\lambda_j\right)\right]$, 1 and $\hat{N}_j$ into the (1-α)CI to get an approximate (1-α)CI for $h\left[\tilde{\rho}_{XY}\left(\lambda_j\right)\right]$. Applying the transformation tanh maps the CI back to [-1, 1] to produce an approximate 100 (1-α)% CI for $\rho_{XY}\left(\lambda_j\right)$ given by

$$\tanh\left\{h\left[\tilde{\rho}_{XY}\left(\lambda_j\right)\right] \pm \zeta_{\frac{\alpha}{2}}\left(\frac{1}{\hat{N}_j - 3}\right)^{1/2}\right\}. \tag{2.11}$$

The quantity $\hat{N}_j$ in Equation (2.11) is the number of the DWT coefficients associated with scale λ_j.

2.3 Application to agitation sedation wavelet modeling

This section presents the application of a wavelets analysis of ICU data of 37 patients. In the hospitalized adult, agitation is defined as excessive verbal behaviour that interferes with patient care, and patient and medical therapies (Chase et al., 2004).

Current agitation management methods rely on subjective agitation assessment, and an appropriate sedation input response, from recorded at bedside agitation scales (Fraser & Riker, 2001b; Jaarsma et al., 2001; Ramsay et al., 1974; Riker et al., 1999; Sessler et al., 2002). The carers then select an appropriate infusion rate based upon their evaluation of these scales, experience, and intuition (Kress et al., 2002). This process is depicted in Figure 2.3 (see Chase et al., 2004).

The recorded infusion rates were obtained from the results reported by Chase et al. (2004). Chase et al. (2004) contains an extensive description of agitation-sedation modelling, as well as other references on agitation-sedation control. The simulated infusion rates are obtained from the equations of Chase et al. (2004). Details of the development and validation of the agitation-sedation (A-S) model can be found in Chase et al. (2004) and see also Rudge et al. (2006a, 2006b).

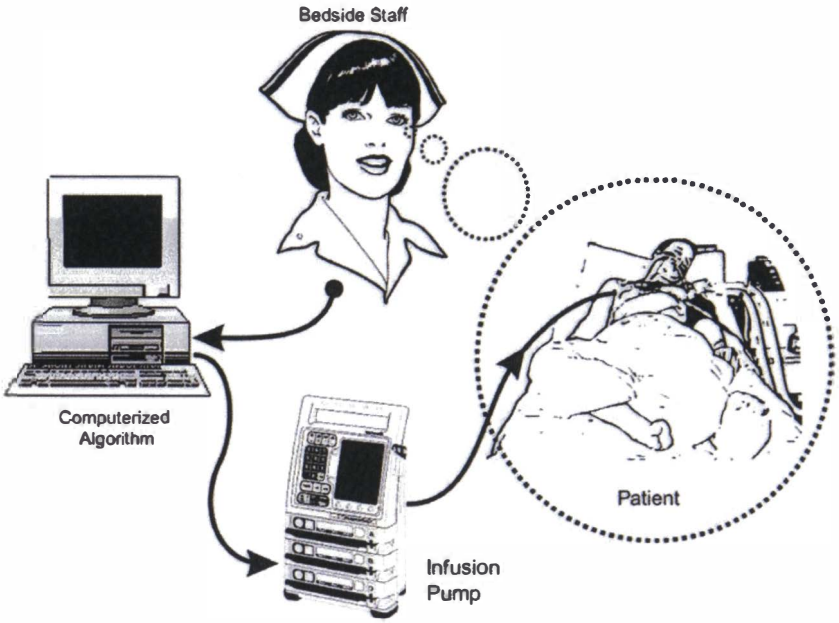


Figure 2.3 Diagram of the feedback loop employing nursing staff feedback of subjectively assessed patient agitation through the infusion controller (diagram is sourced from Chase et al. (2004)).

Figure 2.4 shows patient specific recorded and simulated (as the thick line) profiles for four patients, by way of illustration. The resulting model by Chase et al. (2004) can thus serve as a platform to develop and test semi-automated sedation management controllers that offer the potential of improved agitation management and reduced length of stay in the ICU.

Table A.1 in the appendix gives the equations used, mathematical methods employed and the aims of the given study and the type of performance indicators derived to establish the validity of the A-S model. Our work provides ***new wavelet based diagnostics*** by which to assess the A-S models.

In this chapter we apply the DWT, MODWT and MRA to the infusion profiles (recorded and simulated) of the 37 ICU patents from the work Chase et al. (2004). We provide graphical assessment tools and derive wavelet-based metrics by which to assess the compatibility (closeness) of the empirical (recorded) data and the mathematically modelled (simulated) A-S profiles. These performance measures are based on the modulus of the wavelet correlation at wavelet scale (λ_1) (level 1) and also on summary measures based on wavelet correlations and wavelet cross-correlations (see Table A.1) and the details in this chapter. The aim of our DWT analysis is to see if our wavelets based diagnostics can assess how well the A-S model captures the underlying dynamics between the recorded and the simulated infusion rates at different horizons via the

DWT; and compare our results with those of Chase et al. (2004), Rudge et al. (2006a, 2006b, 2005), and Lee et al. (2005).

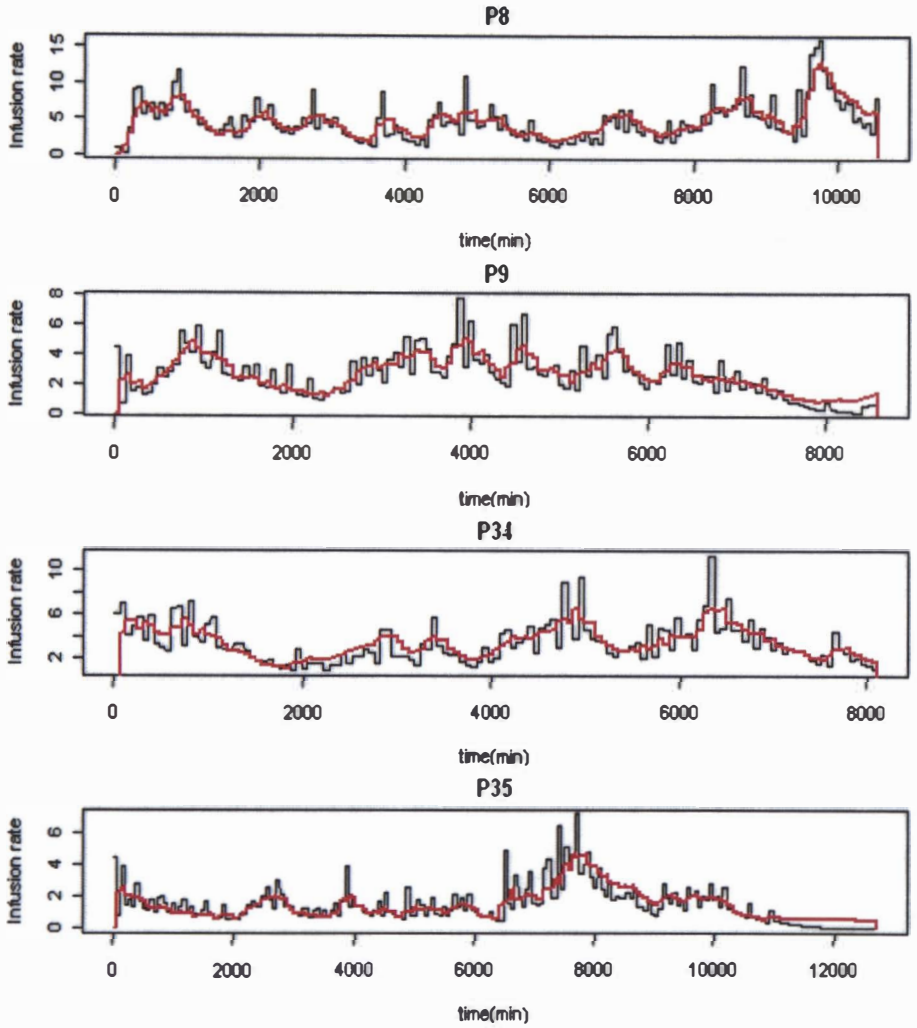


Figure 2.4 Example of the delay between the recorded and simulated (thick line) infusion rate profiles for Patients 8, 9, 34 and 35 (denoted by P8, P9, P34 and P35).

2.3.1 Using the DWT and MODWT

In Section 2.2, we overviewed the DWT and MODWT (Table 2.1). In Section 2.3 we use the DWT, MODWT and their multiresolution analysis (MRA) to examine the "closeness" or similarity between a given patient's recorded and simulated A-S profiles. If any patient shows a similar profile between their recorded infusion rate and simulated infusion rate, we deem it to be a so-called "good tracker", otherwise it is called a poor tracker. We aim to evaluate how well the DWT methods capture these dynamics between the recorded and simulated infusion rate at different time horizons, per patient. Later we develop specific performance indicators to quantify

the 'closeness' between a patient's recorded (R) and simulated (S) profile based on wavelet diagnostics and thereby identify the good versus poor trackers (see Sections 2.3.1-2.3.3).

Figure 2.5 displays the PDWT using the Haar wavelet filter for Patient 2's recorded infusion rate series. Figure 2.6 displays the PDWT (using the Haar wavelet filter) for Patient 2's stimulated infusion rate series. The recorded infusion series is plotted in the upper row of Figure 2.5. There is an obvious increase in the variance of the infusion rate in the first part and the latter third of the series. The length of Patient 2's recorded and simulated series is N=6000 (after sampling) which is divisible by 2^4=16. We can therefore perform an order J_p=4 partial DWT (PDWT) on the R and S series. The first scale of wavelet coefficients, w_1, filters out the high-frequency fluctuations by essentially taking adjacent differences in the data. These series show changes in variance. Comparing the two A-S series (R and S) for Patient 2, there is an obvious decrease in variance in their stimulated infusion series but not in their recorded series. In Figure 2.5, there is a rapid fluctuation in the first part of the recorded series and then rapidly increasing variance in the last part of the recorded series. By contrast P2's simulated infusion series in Figure 2.6 shows fluctuations in the first part, for all wavelet coefficients from w_1 to w_4. Since Patient 2's simulated infusion series does not exhibit low-frequency oscillations, the higher scales with low frequency, namely the wavelet coefficients w_3 and w_4, do not indicate a large variation from zero.

Figure 2.7 displays the sum of squared PDWT coefficients for P2's recorded and simulated infusion series. The distribution of variance will vary depending on the wavelet filter. As the length of the wavelet filter increases, the approximation to an ideal band-pass filter improves and the wavelet filter will better capture the variability in the frequency intervals associated with the DWT wavelet coefficients (Percival & Walden, 2000; Gencay et al., 2001). The sum of squared wavelet coefficients for the Haar wavelet filter for the first two scales then peaks for the vectors w_3 and w_4 of the two series. Patient 2's simulated series $\|w_j\|^2$ varies a small amount between the first three scales, but their recorded series shows a large variation between the first three scales and w_4 (Figure 2.7).

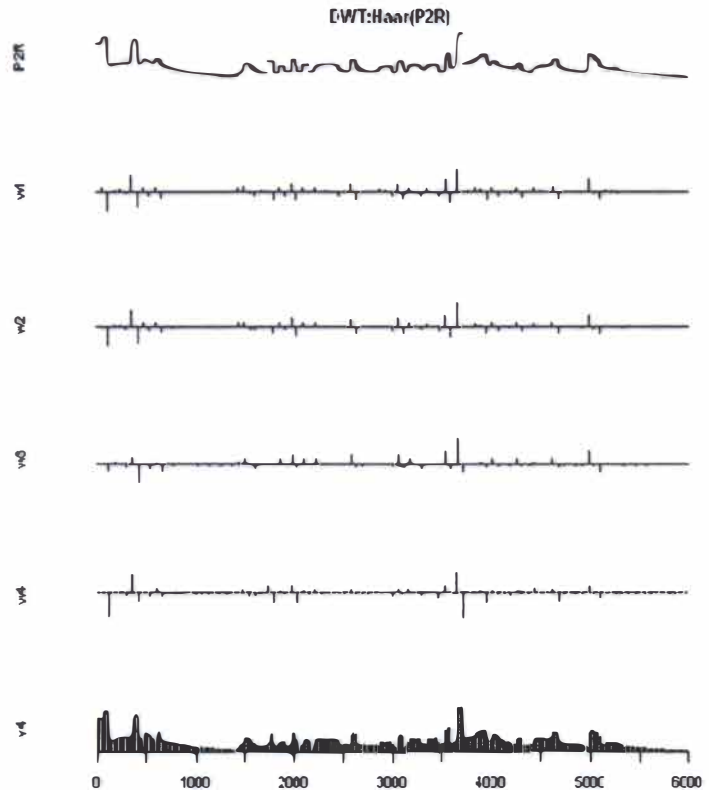


Figure 2.5 Wavelet decomposition of the recorded (R) infusion data for Patient 2 (P2). The upper plot is the original series, the PDWT coefficients vectors are given in the second to the fifth row, and the scaling coefficient vector v4, is last.

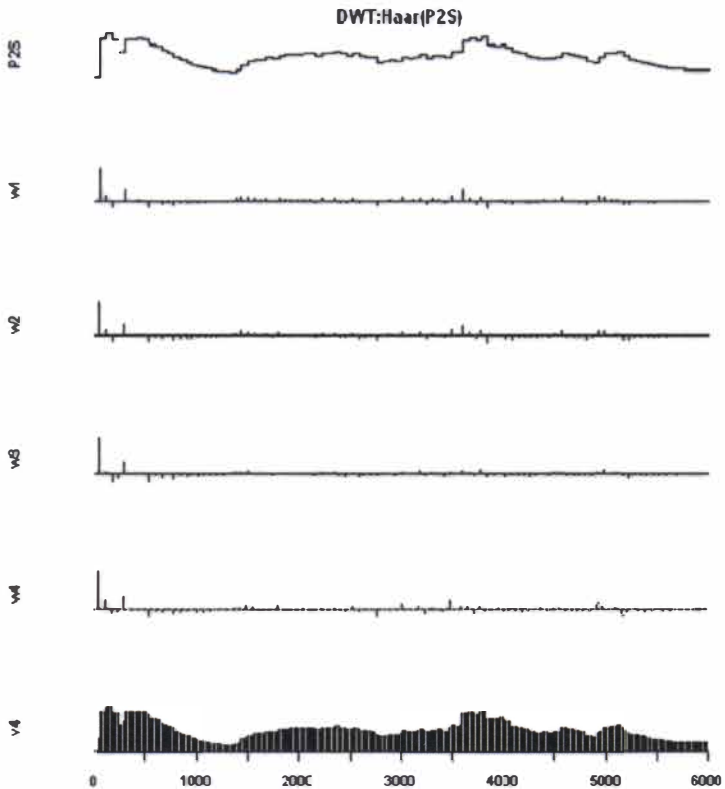


Figure 2.6 Wavelet decomposition of the Patient 2's simulated infusion data. The upper plot is the original series, the PDWT coefficients vectors are shown in the second row to the fifth, and the scaling coefficient vector v4, is given in the last row.

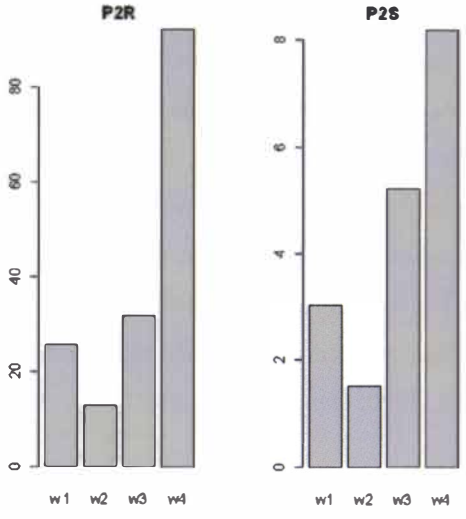


Figure 2.7 Sum of squared wavelet coefficients vectors $\|w_j\|^2, j = 1,...,4,$ for Patient 2's recorded (R) and simulated (S) infusion series, on the LHS and RHS, respectively (using the Haar wavelet filter).

DWT-MRA analysis is formed from four wavelet details and a wavelet smooth, resulting in, $\|X\|^2 = \sum_{j=1}^{J} \|\boldsymbol{d}_j\|^2 + \|\boldsymbol{s}_J\|^2, j = 1,...,4$ (Percival & Walden, 2000; Gencay et al., 2001). DWT-MRA is not much different from the DWT, since if W_j contains a large coefficient, then the wavelet details will contain a large copy of the basis function at that location. This follows from the definitions of $\boldsymbol{d}_j$ and $\boldsymbol{s}_J$ (Percival & Walden, 2000; Gencay et al., 2001). For example, we know that $\boldsymbol{d}_j = \mathcal{W}_j^T \boldsymbol{w}_j$ for $j = 1, \ldots, J$, so the j^{th} level wavelet detail associated with changes in X at scale X_j, is a projection of the j^{th} level wavelet basis vectors onto the level j wavelet coefficients. The smooth $\boldsymbol{s}_J = \sum_{k=J+1}^{J+1} \boldsymbol{d}_k$ defines the jth level wavelet smooth for $0 \le j \le J$. The vector of observations can then be decomposed through the wavelet smooth and rough via $\boldsymbol{x}_J = \boldsymbol{s}_J + \boldsymbol{r}_J$ where $\boldsymbol{r}_J = \sum_{j=1}^{J} \boldsymbol{d}_k$ and $1 \le j \le J+1$. The DWT possesses key attributes like approximately de-correlating certain processes and efficiently capturing important features of a process via limited number of coefficients (therefore, being data reductive). However there are also limitations in applying the DWT to time series. The DWT is very useful, but does not posses all attributes that may be desirable for certain applications. In response to this, an alternative wavelet transform has been developed called the MODWT.

Figure 2.8 displays the MODWT-MRA of Patient 2's recorded infusion series (using LaDaub (8)), which is the least asymmetric Daubechies wavelet filter. The MODWT is no longer limited to decomposing a sample size of dyadic length. We can thus take J=9 since the only limiting factor

is the overall depth of the transform given by $J = \log_2 6000 = 12.55$ for Patient 2. When we compare Figure 2.6 against Figure 2.8, there do not appear to be any obvious differences due to the different wavelet basis functions chosen. The correlation between coefficients is effectively smoothing over features specific to the wavelet filter. Hence, even when using non-smoothing wavelet filters, such as the Haar wavelet filters, we may observe reasonably smooth wavelet details and smooth. A crucial characteristic to an MRA, which employs the MODWT, is that the wavelet details and smooth are related to zero-phase filters (see Table 2.1). Thus characteristics of note in the wavelet details and smooth can be entirely aligned with the original series under examination (Percival & Walden., 2000; Gencay et al., 2001).

This latter feature is not available through the DWT since it subsamples the output of its filtering operations. Thus, we observe the following increased recorded infusion rate in the wavelet details in the first peak around 60 minutes and third part around 3660 minutes as a function of scale in Figure 2.8. Figure 2.9 shows a rapidly increased magnitude in d_1, d_3, d_5, and d_7 at observations around 61-66 minutes, but shows a small variation around zero in the third part of the detail coefficients. The two wavelet smooth profiles (s9) in Figure 2.8 and 2.9 exhibit a similar profile but do not give detailed information about the variation. Therefore we can conclude that the recorded infusion series and simulated infusion series for Patient 2 differ. Indeed Patient 2 is confirmed to be a so-called "poor tracker" via wavelet diagnostics proposed here (see Section 2.2.1 and 2.2.2).

The distortion as like the smooth series turn upwards which are the bottom series in Figure 2.8 and 2.9 even the time series is almost level due to boundary conditions. A method for handling wavelet transform boundary condition that is appropriate for the simulated and recorded infusion series whose behavior at the boundaries is roughly

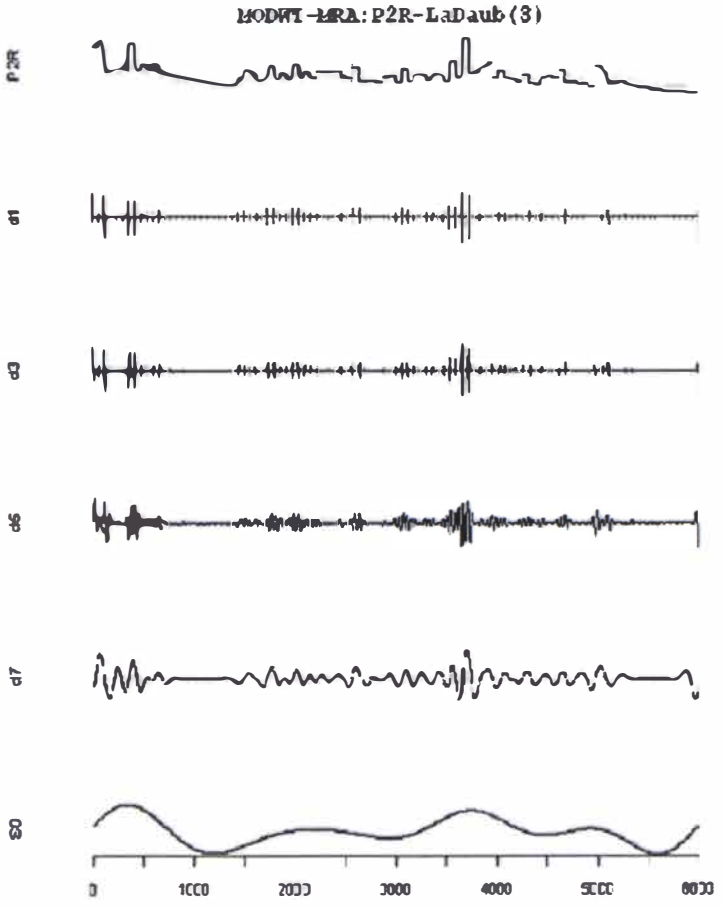


Figure 2.8 LaDaub (8) MODWT multiresolution analysis (MRA) of the recorded infusion series of Patient 2 (P2R).

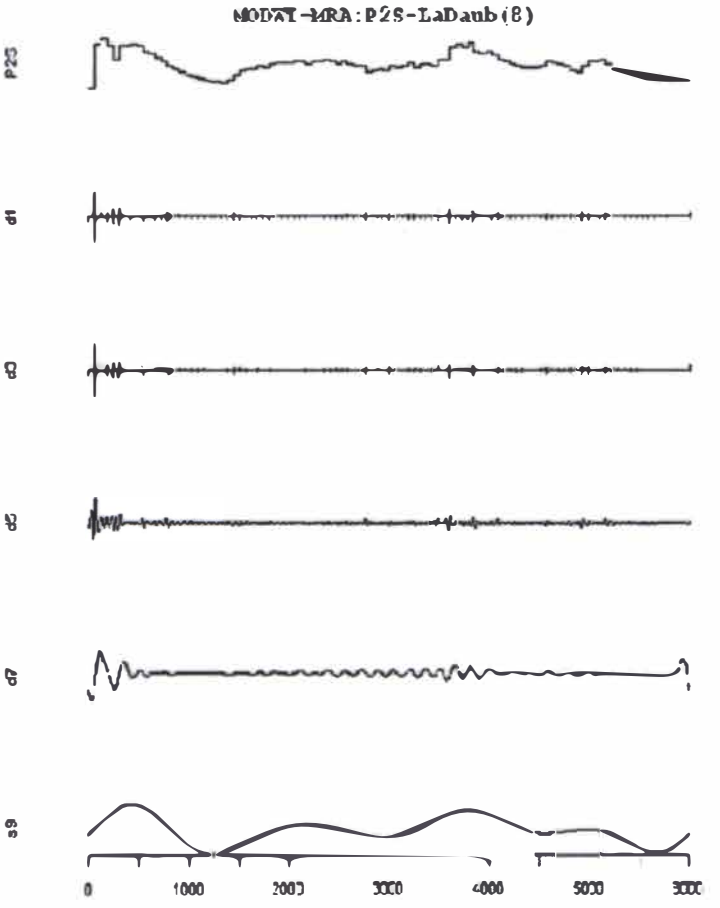


Figure 2.9 LaDaub (8) MODWT multiresolution analysis (MRA) of the simulated infusion series of Patient 2 (P2S).

characterized by a linear increase or decrease. Due to the large difference between the beginning and end of the simulated and recorded infusion series, the MODWT was computed using reflection boundary conditions rather than the periodic conditions. If the time series show very strong seasonality, the periodic conditions applied. A common technique is to reflect the time series about its last observation since the reflecting the series does not change the sample mean nor the sample variance.

Two series the bottom series in Figure 2.8 and 2.9 appear similar but not actually identical. Two smooth series is plotted in Figure 2.9.1 look quite similar presumably their cross-correlation would be close to unity that they have a similar pattern. The smooth is a quite useful estimate of trend in time series.

The lag/lead relationship between the simulated and recorded infusion series, however, can be investigated on a scale-by-scale basis via a MODWT-MRA (using the LaDaub (8) filter). The results of a MODWT-MRA are shown in Table 2.3. Fourteen patients (patients 3, 9, 11, 17, 20, 22, 26, 30, 33, 34, 35, 36, 37) have recorded infusion series that lead the simulated infusion series (as indicated by a bolded time lag in Table 2.3).

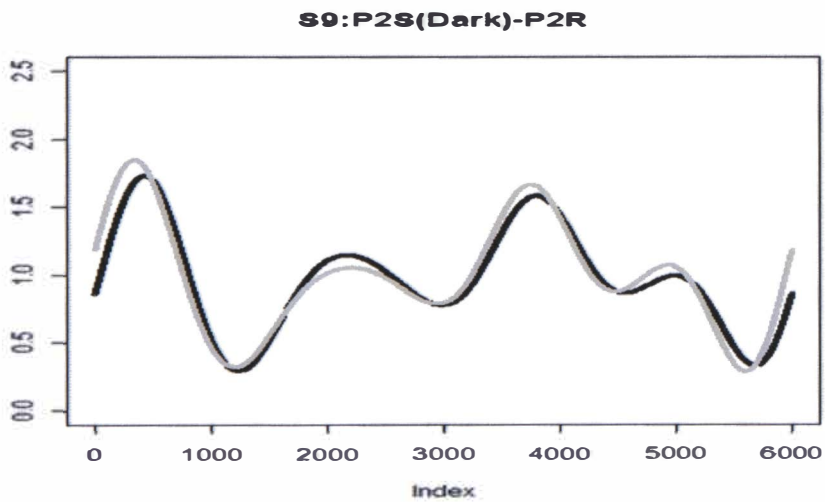


Figure 2.9.1 Comparison two smooth series between the simulated and recorded infusion series using LaDaub (8) MODWT-MRA of Patient 2.

Table 2.3 The maximum wavelet coefficient location of the MODWT-MRA (using LaDaub (8)). R-recorded infusion series and S-simulated infusion series

Patient no.	Series type	Wavelet details (scales)					Average (d1,…,d5)	Time lag (min)
		d1	d2	d3	d4	d5		
P1	R	60	59	58	55	51	57	8
	S	61	62	63	66	71	65	(=65-57)
P2	R	1	2	3	6	11	5	60
	S	61	62	63	66	71	65	(=65-5)
P3	R	1	2	3	6	10	4	61
	S	61	62	63	66	71	65	(=65-4)
P4	R	2101	2102	2103	2106	2110	2104	**59**
	S	2041	2042	2043	2046	2051	2045	(=2104-2045)
P5	R	2341	2342	2343	2346	2350	2344	592
	S	2940	2939	2938	2935	2930	2936	(2935-2344)
P6	R	5341	5342	5343	5346	5350	5344	352
	S	5700	5699	5698	5695	5690	5696	(=5696-5344)
P7	R	1	2	3	6	11	5	60
	S	61	62	63	66	71	65	(=65-5)
P8	R	4801	4802	4803	4806	4810	4804	5752
	S	10560	10559	10558	10555	10550	10556	(=10556-4804)
P9	R	3900	3899	3898	3895	3891	3897	**3832**
	S	61	62	63	66	71	65	(=3897-65)
P10	R	14761	14762	14763	14766	14770	14764	5932
	S	20700	20699	20698	20695	20690	20696	(=20696-14764)
P11	R	3000	2999	2998	2995	2991	2997	**352**
	S	2641	2642	2643	2646	2651	2645	(=2997-2645)
P12	R	1	2	3	6	10	4	61
	S	61	62	63	66	71	65	(=65-4)
P13	R	3001	3002	3003	3006	3010	3004	2152
	S	5160	5159	5158	5155	5150	5156	(=5156-3004)
P14	R	1	2	3	6	10	4	61
	S	61	62	63	66	71	65	(=65-4)
P15	R	4981	3302	3303	3306	3310	3640	1336
	S	4980	4979	4978	4975	4970	4976	(=4976-3640)
P16	R	7080	7079	7078	7075	7071	7077	6540
	S	13620	13619	13618	13615	13610	13616	(=13616-7077
P17	R	121	122	123	126	130	124	**59**
	S	61	62	63	66	71	65	(=124-65)
P18	R	3181	3182	3183	3186	3190	3184	173
	S	3360	3359	3358	3355	3351	3357	(=3357-3184)
P19	R	5881	5882	5883	5886	5890	5884	2032
	S	7920	7919	7918	7915	7910	7916	(=7916-5884)
P20	R	1741	1742	1743	1746	1750	1744	**59**
	S	1681	1682	1683	1686	1691	1685	(=1744-1685)
P21	R	960	959	958	955	951	957	2188
	S	3720	3719	3718	3715	851	3145	(=3145-957)
P22	R	5581	5582	5583	5586	5590	5584	**5519**
	S	61	62	63	66	71	65	(=5584-65)
P23	R	2701	2702	2703	2706	2711	2705	**60**
	S	2641	2642	2643	2646	2651	2645	(=2705-2645)
P24	R	7920	7919	7918	7915	7911	7917	539
	S	8460	8459	8458	8455	8450	8456	(=8456-7917)
P25	R	3061	3062	3063	3066	3070	3064	772
	S	3840	3839	3838	3835	3830	3836	(=3836-3064)
P26	R	121	122	123	126	130	124	**59**
	S	61	62	63	66	71	65	(=124-65)

P27	R	1	2	3	6	10	4	61
	S	61	62	63	66	71	65	(=65-4)
P28	R	1621	1622	1623	1626	1630	1624	10612
	S	12240	12239	12238	12235	12230	12236	(=12236-1624)
P29	R	1	2	3	6	10	4	61
	S	61	62	63	66	71	65	(=65-4)
P30	R	2761	2762	2763	2766	2770	2764	**59**
	S	2701	2702	2703	2706	2711	2705	(=2764-2705)
P31	R	1981	1982	1983	1986	1990	1984	1919
	S	61	62	63	66	71	65	(=1984-65)
P32	R	6661	6662	6663	6666	6671	6665	8511
	S	15180	15179	15178	15175	15170	15176	(=15176-6665)
P33	R	4681	4682	4683	4686	4690	4684	**59**
	S	4621	4622	4623	4626	4631	4625	(=4684-4625)
P34	R	6360	6359	6358	6355	6351	6357	**6292**
	S	61	62	63	66	71	65	(=6357-65)
P35	R	7381	7382	7383	7386	7390	7384	7319
	S	61	62	63	66	71	65	(=7384-65)
P36	R	541	542	543	546	11	437	372
	S	61	62	63	66	71	65	(=437-65)
P37	R	2461	2462	2463	2466	2470	2464	59
	S	2401	2402	2403	2406	2411	2405	

It is noteworthy that of the patients exhibiting a lagged discrepancy, our wavelet diagnostics and those of Chase et al. (2004) and Rudge et al. (2006b) identify the following poor performers in common (Patients 9, 11, 17, 22, 33, 34 and 35). The simulated profile peaks later than the patient's recorded infusion possibly due to the delay in distribution time for the drug. This result implies that, while performing well most of the time, the simulated rate deviates from the recorded infusion rate over particular periods, and takes some time before tending towards the infusion rate again.

These periods indicate times of the patient's hospital length of stay in ICU, where the model may not capture the subject's specific A-S dynamics (see the time lags as such given in the last column of Table 2.3). These periods may correspond to periods of marked distress or physiological alterations due to the patient's state. In particular, a common reason for the departure of the simulated profile (Table 2.3) is this apparent time-lag. Particularly small departures indicate rapid increases (or decreases) in the recorded infusion rate, where the simulated infusion rate appears to lag behind. These differences may be a result of the medical staff's over or under-assessment of the patient's agitation.

Although DWT has several appealing qualities that make it a useful method for time series analysis, exhibiting features that vary in both time and frequency, DWT is not optimal in capturing the underlying dynamics between the recorded and the simulated infusion rate for the

ICU patients. In Section 2.3.2-2.3.3 we analyse the association between the recorded and the simulated infusion profiles per patient using wavelet correlation, wavelet variance, and wavelet cross-correlation diagnostics based on a DWT approach.

2.3.2 Using the Wavelet Variance and Correlation

An important characteristic of the DWT and MODWT is their ability to decompose the variance of a stochastic process. They provide a decomposition of the variance between the original series and the DWT (or MODWT) wavelet coefficients (Percival & Walden, 2000;Gencay et al., 2001). As mentioned before, a vector of wavelet coefficients is associated with changes at a particular wavelet scale. Applying the DWT or MODWT to a stochastic process thus produces decomposition on a scale-by- scale basis (Percival & Walden, 2000; Gencay et al., 2001).

We now decompose the variance of both the recorded and simulated infusion profiles of the 37 patients using the MODWT. Figure 2.10 displays the MODWT-based wavelet variance, using the Haar wavelet filter, of a selected group of patients. There are virtually no differences between the recorded and the simulated infusion rates in Figure 2.10 except at the last scale (of 128). This means that the two series for each patient appear to exhibit the same variability at all other remaining seven wavelet scales (Figure 2.10).

We now investigate how well the recorded and simulated series are associated with one another. The wavelet covariance is a measure of this association when comparing the two series at the same time point. Figure 2.11 provides the MODWT-based covariance (using the Daub (4) wavelet filter) for Patients 2, 4, 8 and 14.

The wavelet covariance (WCOV) follows a similar pattern to the wavelet variance (see Figure 2.11). It shows that the WCOV is very close to zero up to the fourth wavelet scale and differences are shown at the eighth wavelet scale (λ_8). Although there appears to be mostly a positive association between the two series, it is difficult to compare the wavelet scales because of the differing variability exhibited by each of them. Standardizing the WCOV by the variance of the given bivariate series, at each scale, is a simple way of overcoming this problem and makes it possible to compare the magnitude of the association across scales.

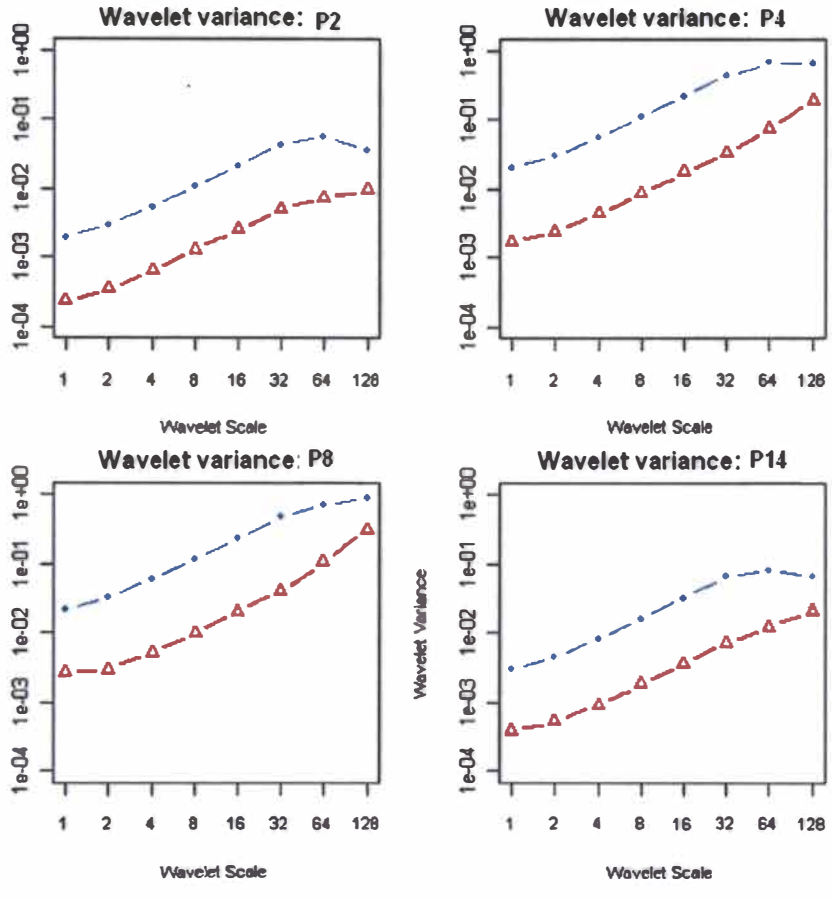


Figure 2.10 Wavelet variances for Patients 2, 4, 8 and 14. The lines with circles represent the recorded infusion rate and the lines with a triangle represent the simulated infusion rate for each patient.

The wavelet correlation (WCORR) is thus used, instead of the wavelet covariance, to investigate how the simulated and recorded infusion series correlate. Recall that the wavelet correlation is the correlation between the scale λ_j wavelet coefficients of the bivariate time series, and is simply the wavelet covariance standardized at each scale (see Equation (2.7)). In this study an estimate of WCORR, $\tilde{\rho}$ from Equation (2.7), is used to investigate the wavelet scale. Differences were shown at the eighth wavelet scale (λ_8).

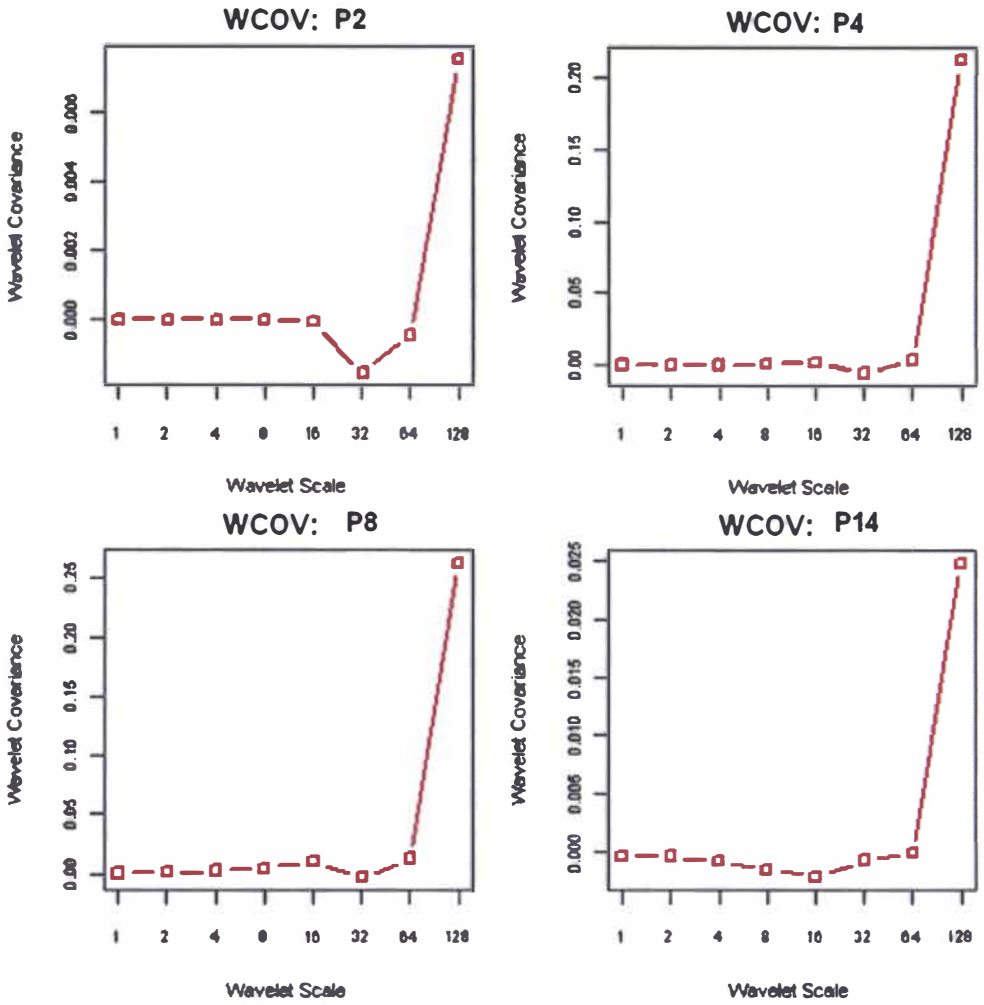


Figure 2.11 Wavelet covariance (WCOV) between the recorded infusion rate and the simulated infusion rate for Patients 2, 4, 8 and 14.

Table 2.4 shows the median of the WCORR values of for the first five wavelet scales and of the full eight wavelet scales; in addition to the AND (average normalized density), RAND (relative average normalized density), and TI (tracking index) values for all 37 ICU patients. The AND and RAND values are from Rudge et al. (2006b), and the TI values were derived by using weighted kernel density estimation by Chase et al (2004). We note that even with increasing the number of wavelet scales the median is not much different from those at the smaller wavelet scale levels (Table 2.4). This gives strong justification to use the PDWT instead of a full DWT. In practice, PDWTs are more useful than the full DWT with the scaling coefficients able to capture the large scale variation with low frequency in the time series under investigation.

For each patient i (i=1, 2,..., 37) we count the number of its wavelet scales (out of a maximum of 8) for which the WCORR value at scale λ_j is not significant (at the 5% level of significance) (see

Table 2.6). This variable is denoted by "Count of NS" and given per patient, with specification of the significance of WCORR at λ_j (j=1, 2, …, 8) as either significant (S) or not significant (NS) in Table 2.5. Patient specific AND (average normalized density), RAND (relative average normalized density), and TI (tracking index) values, all derived by Rudge et al (2006b) and Chase et al. (2004) are also shown in Table 2.5. Definitions for the AND, RAND and TI performance indicators or diagnostics are detailed in Table A.1. Poor trackers are identified now via wavelet diagnostics as follows: Patients with a "Count of NS" greater or equal to 2 and a non-significant WCORR value at scale λ_1 (level 1) are said to be a poor tracker; as are patients with a "Count of NS" less or equal to 3 and a *significant negative* WCORR value at scale λ_1 (level1) and a *significant negative* WCORR value at scale λ_6 (level6). Table 2.5 indicates in bold such 15 poor trackers as defined by our our diagnostics ("Count of NS" and WCORR at scales 1 and 6), derived in this Chapter. This means that their simulated A-S profiles do not mirror their actual (recorded) A-S infusion profile closely over time according to the wavelet correlational values. Alternative wavelet probability band diagnostics are derived in Chapter 3 to further quantify how well a patient tracks.

Figure 2.12 shows the estimated WCORR, $\tilde{\rho}$ and 95% CI for four patients, P2 and P4 (poor trackers) and P8 and P14 (good trackers). From Figure 2.12 we see that WCORR is generally significant for wavelet scales 1, 2, 4, 8 and 16 for so-called good trackers (whether a positive or negative WCORR).

This is not the case for so-called poor trackers, who generally exhibit non-significant WCORR at λ_j for $j \leq 5$.

As mentioned above Table 2.5 also lists patients which are considered to track poorly according to both Rudge et al's (2006b) and Chase et al's (2004) diagnostics, which are based on completely different methodologies as summarised in Table A.1. It is noteworthy that 11 of our so called 13 poor trackers are also considered to be "poor trackers" by either or both of Rudge et al's (2006b) and Chase et al's (2004) performance indicates. By using the data per patient from Table 2.5, we can perform a Kruskal Wallis test to statistically compare both the performance indicators of the good trackers and the so-called poor trackers. These results are summarized in Table 2.6 and Table 2.7.

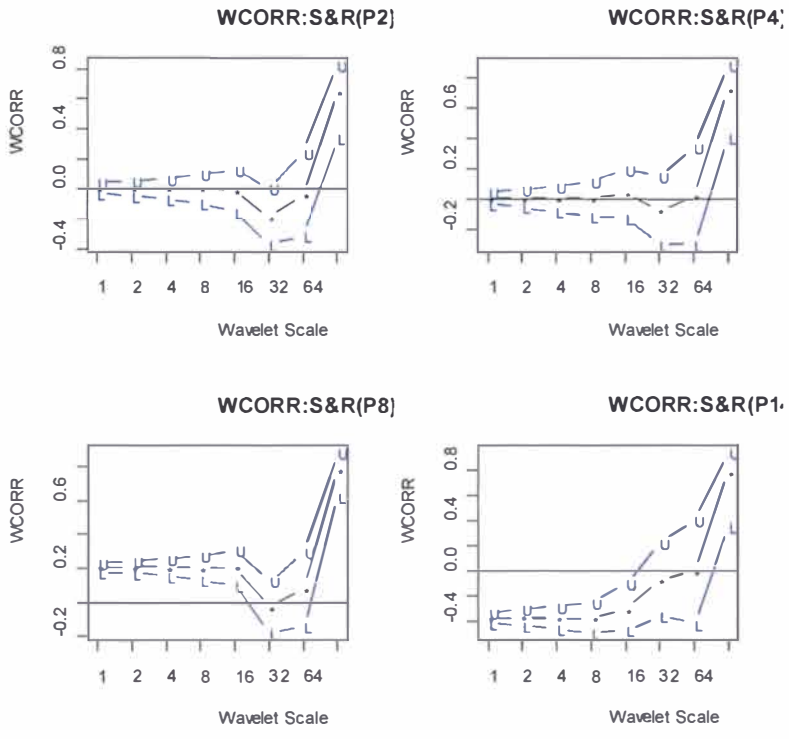


Figure 2.12 Wavelet correlation for Patient 2 (Top LHS), Patient 4 (Top RHS), Patient 8 (bottom LHS) and Patient 14 (bottom RHS) with the approximate 95% confidence interval. Patients 2 and 4 are poor trackers in contrast to Patients 8 and 14 who are good trackers with >4 significant WCORR at wavelet scales (1, 2, 4, 8,and 16).

Table 2.4 Median of Wavelet Correlation between the recorded and simulated series, in addition to the AND, RAND, and TI measures for each of the patients.

Patient no.	Median & 95% CI of 5 λ_j	Median & 95% CI of 8 λ_j	AND	RAND	TI (SE)
P1	-0.134 (-0.146, 0.122)	-0.114 (-0.305, 0.078)	0.51	0.62	87.0 (0.041)
P2	0.005 (0.000,0.011)	0.005 (-0.167, 0.178)	0.53	0.66	86.2 (0.03)
P3	-0.504 (-0.529, 0.479)	-0.471 (-0.751, -0.191)	0.70	0.83	88.8 (0.015)
P4	0.006 (0.006,0.018)	0.006 (-0.173, 0.185)	0.56	0.62	80.1 (0.09)
P5	0.073 (0.071, 0.075)	0.073 (-0.098, 0.244)	0.60	0.80	91.1 (0.016)
P6	-0.202 (-0.220, 0.184)	-0.185 (-0.392, 0.023)	0.70	0.84	87.0 (0.014)
P7	-0.081 (-0.098, 0.064)	-0.074 (-0.281, 0.134)	0.33	0.43	84.5 (0.06)
P8	0.206 (0.204, 0.208)	0.206 (0.041, 0.371)	0.45	0.59	87.4 (0.027)
P9	-0.021 (-0.034, -0.008)	-0.021 (-0.220, 0.179)	0.49	0.62	87.3 (0.02)
P10	0.051 (0.051, 0.052)	0.051 (-0.097, 0.199)	0.27	0.34	83.4 (0.04)
P11	-0.001 (-0.009, 0.007)	0.001 (-0.157, 0.158)	0.31	0.38	83.7 (0.08)
P12	-0.359 (-0.376, 0.342)	-0.338 (-0.577, -0.098)	0.61	0.77	84.1 (0.033)
P13	0.258 (0.258, 0.258)	0.258 (0.104, 0.412)	0.37	0.45	86.1 (0.072)
P14	-0.379 (-0.410, 0.348)	-0.340 (-0.618, -0.061)	0.48	0.56	93.1 (0.034)
P15	0.092 (0.088, 0.096)	0.085 (-0.105, 0.275)	0.45	0.60	91.1 (0.011)
P16	0.169 (0.167, 0.171)	0.169 (0.000, 0.338)	0.44	0.57	87.9 (0.021)
P17	-0.122 (-0.143, 0.101)	-0.122 (-0.302, 0.058)	0.61	0.72	84.0 (0.037)
P18	-0.134 (-0.136, 0.132)	-0.134 (-0.329, 0.061)	0.55	0.68	94.6 (0.026)
P19	0.273 (0.271, 0.275)	0.273 (0.142, 0.404)	0.50	0.66	91.1 (0.014)
P20	0.057 (0.057, 0.057)	0.057 (-0.095, 0.209)	0.53	0.65	87.3 (0.033)
P21	0.017(0.013, 0.021)	0.017 (-0.181, 0.215)	0.53	0.72	78.5 (0.09)
P22	-0.051(-0.060, -0.042)	-0.051 (-0.247, 0.145)	0.35	0.45	85.2 (0.04)
P23	0.149 (0.143, 0.155)	0.149 (-0.006, 0.304)	0.72	0.85	84.8 (0.105)
P24	0.232 (0.231, 0.232)	0.232 (0.073, 0.391)	0.43	0.54	88.1 (0.023)
P25	-0.214 (-0.224, 0.204)	-0.213 (-0.401, -0.024)	0.50	0.66	92.4 (0.025)
P26	-0.123 (-0.143, 0.103)	-0.098 (-0.264, 0.069)	0.68	0.88	87.4 (0.031)
P27	-0.108 (-0.116, -0.100)	-0.108 (-0.306, 0.090)	0.39	0.49	74.9 (0.07)
P28	-0.101 (-0.105, -0.097)	-0.101 (-0.286, 0.084)	0.34	0.44	89.2 (0.02)
P29	-0.616 (-0.629, 0.603)	-0.599 (-0.897, -0.301)	0.38	0.45	77.6 (0.083)
P30	-0.044 (-0.046, 0.042)	-0.045 (-0.229, 0.140)	0.63	0.82	92.2 (0.021)
P31	0.040 (0.037, 0.043)	0.040 (-0.131, 0.211)	0.40	0.51	89.3 (0.030)
P32	0.035 (0.032, 0.038)	0.035 (-0.151, 0.220)	0.38	0.50	89.3 (0.02)
P33	0.046 (0.045, 0.047)	0.046 (-0.125, 0.217)	0.28	0.36	88.7 (0.020)
P34	-0.019 (-0.023, -0.015)	-0.019 (-0.207, 0.169)	0.43	0.55	86.5 (0.03)
P35	-0.172 (-0.185, 0.159)	-0.169 (-0.350, 0.012)	0.38	0.46	85.9 (0.044)
P36	0.081 (0.071, 0.091)	0.081 (-0.095, 0.257)	0.52	0.64	86.4 (0.095)
P37	0.273 (0.272, 0.273)	0.273 (0.142, 0.404)	0.53	0.59	79.9 (0.093)

Table 2.5 Wavelet correlation analysis of the 37 ICU patients. S-significant, NS-non-significant WCORR at given scale λ_j (j= 1, 2, …, 8). Bolded patients indicate poor trackers according to the WCORR and "Count NS" values.

Patient ID	Wavelet scale λ_j (j= 1, 2, …, 8)								Count of NS*	Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	1	2	4	8	16	32	64	128		AND▲	RAND▲	TI (SE)
P1	S	S	S	S	S	NS	NS	S	2	0.51	0.62	87.0 (0.041)
P2+	NS	NS	NS	NS	NS	NS	NS	S	7	0.53	0.66	86.2 (0.037)
P3	S	S	S	S	S	S	N.S	S	1	0.70	0.83	88.8 (0.015)
P4	NS	NS	NS	NS	NS	NS	NS	S	7	0.56	0.62	80.1 (0.095)
P5	S	S	NS	NS	NS	NS	NS	S	5	0.60	0.80	91.1 (0.016)
P6	S	S	S	S	S	NS	NS	S	2	0.70	0.84	87.0 (0.014)
P7	S	S	NS	NS	NS	NS	NS	S	5	0.33	0.43	84.5 (0.068)
P8	S	S	S	S	S	NS	NS	S	2	0.45	0.59	87.4 (0.027)
P9	NS	NS	NS	NS	NS	NS	NS	S	7	0.49	0.62	87.3 (0.024)
P10	S	S	S	NS	NS	S	NS	S	3	0.27	0.34	83.4 (0.041)
P11	NS	NS	NS	NS	NS	NS	NS	S	7	0.31	0.38	83.7 (0.080)
P12	S	S	S	S	S	S	NS	S	1	0.61	0.77	84.1 (0.033)
P13	S	S	S	S	S	NS	NS	S	2	0.37	0.45	86.1 (0.072)
P14	S	S	S	NS	S	NS	NS	S	3	0.48	0.56	93.1 (0.034)
P15	S	S	S	NS	NS	NS	NS	S	4	0.45	0.60	91.1 (0.011)
P16	S	S	S	S	S	NS	NS	S	2	0.44	0.57	87.9 (0.021)
P17	S	S	S	S	S	NS	NS	S	2	0.61	0.72	84.0 (0.037)
P18	S	S	S	S	NS	S	NS	S	2	0.55	0.68	94.6 (0.026)
P19	S	S	S	S	S	NS	NS	S	2	0.50	0.66	91.1 (0.014)
P20	S	S	S	NS	NS	NS	NS	S	4	0.53	0.65	87.3 (0.033)
P21	NS	NS	NS	NS	NS	NS	NS	S	7	0.53	0.72	78.5 (0.095)
P22	S	S	NS	NS	NS	NS	NS	S	5	0.35	0.45	85.2 (0.043)
P23	S	S	S	S	NS	NS	NS	S	3	0.72	0.85	84.8 (0.105)
P24	S	S	S	S	S	NS	NS	S	2	0.43	0.54	88.1 (0.023)

P25	S	S	S	S	S	NS	NS	S	2	0.50	0.66	92.4 (0.025)
P26	S	S	S	NS	NS	NS	NS	S	4	0.68	0.88	87.4 (0.031)
P27	S	S	S	S	NS	S	S	S	1	0.39	0.49	74.9 (0.074)
P28	S	S	S	S	NS	S	NS	S	2	0.34	0.44	89.2 (0.027)
P29	S	S	S	S	S	S	NS	S	1	0.38	0.45	77.6 (0.083)
P30	NS	NS	NS	NS	NS	NS	S	S	6[Ω]	0.63	0.82	92.2 (0.021)
P31	S	S	NS	NS	NS	NS	NS	S	5	0.40	0.51	89.3 (0.030)
P32	S	S	NS	NS	NS	S	NS	S	4	0.38	0.50	89.3 (0.022)
P33	S	S	S	NS	NS	S	NS	S	3	0.28	0.36	88.7 (0.020)
P34	NS	NS	NS	NS	NS	S	NS	S	6	0.43	0.55	86.5 (0.034)
P35	S	S	S	S	S	S	NS	S	1	0.38	0.46	85.9 (0.044)
P36	S	S	NS	NS	NS	NS	NS	S	5	0.52	0.64	86.4 (0.095)
P37	S	S	S	S	S	NS	NS	S	2	0.53	0.59	79.9 (0.093)

+ bolded indicates a patient who is a poor tracker, i.e. their simulated infusion profiles is not close nor well correlated with their actual recorded infusion profile.

* Count of NS gives the number of nonsignificant (NS) wavelet correlations at scales λ_j (j=1, 2, …, 8). The maximum value of NS is 8.0.

▲ AND, RAND and TI diagnostics are defined in Table A.1.

Ω Whilst Patient 30 has NS=6, which is relatively large, its high AND=0.63, RAND=0.82, and TI=92.2% are indicative of good tracking.

Table 2.6 Kruskal Wallis test for the wavelet correlation, Rudge et al's (2006b) and Chase et al's (2004) diagnostics for the wavelet based poor versus good tracker groups.

Poor Pt ID	Wavelet Correlations (scale λ_j)								Rudge et al. (2006b) diagnostics		Chase et al. (2004) diagnostics
	□□	□□	□□	□□	□□	□□	□□	□□	AND	RAND	TI(SE)
P2	0.005	0.005	0.005	0.005	-0.008	-0.186	-0.034	0.649	0.53	0.66	86.2 (0.03)
P4	0.006	0.006	0.005	0.005	0.036	-0.077	0.023	0.724	0.56	0.62	80.1 (0.09)
P7	-0.081	-0.081	-0.081	-0.081	-0.037	-0.066	0.034	0.785	0.33	0.43	84.5 (0.06)
P9	-0.020	-0.021	-0.021	-0.021	0.012	-0.066	0.012	0.793	0.49	0.62	87.3 (0.02)
P10	0.051	0.051	0.051	0.051	0.052	-0.125	-0.040	0.591	0.27	0.34	83.4 (0.04)
P11	-0.001	-0.001	-0.001	-0.001	0.019	0.093	0.002	0.654	0.31	0.38	83.7 (0.08)
P21	0.016	0.017	0.017	0.017	0.026	-0.148	-0.031	0.779	0.53	0.72	78.5 (0.09)
P22	-0.051	-0.051	-0.051	-0.051	-0.027	-0.115	-0.052	0.739	0.35	0.45	85.2 (0.04)
P27	-0.108	-0.108	-0.108	-0.108	-0.088	-0.227	-0.201	0.660	0.39	0.49	74.9 (0.07)
P28	-0.100	-0.101	-0.101	-0.101	-0.090	-0.200	-0.157	0.628	0.34	0.44	89.2 (0.02)
P29	-0.616	-0.616	-0.616	-0.616	-0.582	-0.497	-0.294	0.627	0.38	0.45	77.6 (0.083)
P32	0.032	0.034	0.035	0.035	0.042	-0.139	-0.061	0.732	0.38	0.50	89.3 (0.02)
P33	0.046	0.046	0.046	0.046	0.049	-0.127	-0.012	0.691	0.28	0.36	88.7 (0.020)
P34	-0.019	-0.019	-0.019	-0.019	-0.029	-0.213	-0.131	0.676	0.43	0.55	86.5 (0.03)
P35	-0.172	-0.172	-0.172	-0.172	-0.139	-0.155	-0.166	0.576	0.38	0.46	85.9 (0.044)
Overall MD (95%CI)	-0.019 (-0.092 ,0.013)	-0.019 (-094, 0.013)	-0.019 (-0.094, 0.013)	-0.019 (-094, 0.013)	-0.008 (-0.069, 0.032)	-0.139 (-0.195, -0.091)	0.040 (-0.147, -0.003)	0.676 (0.636, 0.736)	0.390 (0.343, 0.515)	0.500 (0.444, 0.620)	85.2 (81.333, 87.001)
Good Pt ID											
P1	-0.112	-0.134	-0.134	-0.134	-0.107	-0.101	-0.115	0.662	0.51	0.62	87.0 (0.041)
P3	-0.503	-0.504	-0.504	-0.504	-0.439	-0.221	-0.242	0.675	0.7	0.83	88.8 (0.015)
P5	0.074	0.073	0.073	0.073	0.078	-0.118	-0.004	0.707	0.6	0.80	91.1 (0.016)
P6	-0.202	-0.202	-0.202	-0.202	-0.155	-0.090	-0.167	0.666	0.7	0.84	87.0 (0.014)
P8	0.208	0.206	0.206	0.206	0.211	-0.029	0.081	0.783	0.45	0.59	87.4 (0.027)
P12	-0.359	-0.359	-0.359	-0.359	-0.316	-0.264	-0.179	0.645	0.61	0.77	84.1 (0.033)
P13	0.258	0.258	0.258	0.258	0.257	0.007	0.107	0.768	0.37	0.45	86.1 (0.072)
P14	-0.378	-0.379	-0.380	-0.380	-0.301	-0.074	-0.003	0.785	0.48	0.56	93.1 (0.034)
P15	0.086	0.092	0.093	0.093	0.084	-0.149	-0.039	-0.697	0.45	0.6	91.1 (0.011)
P16	0.168	0.169	0.169	0.169	0.173	-0.067	0.035	0.758	0.44	0.57	87.9 (0.021)
P17	-0.122	-0.122	-0.122	-0.122	-0.069	0.131	-0.161	0.604	0.61	0.72	84.0 (0.037)
P18	-0.134	-0.134	-0.134	-0.134	-0.130	-0.243	-0.211	0.628	0.55	0.68	94.6 (0.026)
P19	0.272	0.273	0.273	0.273	0.277	0.066	0.195	0.726	0.5	0.66	91.1 (0.014)
P20	0.057	0.057	0.057	0.057	0.057	-0.106	-0.051	0.613	0.53	0.65	87.3 (0.033)
P23	0.149	0.149	0.149	0.149	0.164	-0.019	0.099	0.729	0.72	0.85	84.8 (0.105)
P24	0.231	0.232	0.232	0.232	0.232	-0.007	0.157	0.793	0.43	0.54	88.1 (0.023)

P25	-0.211	-0.214	-0.214	-0.214	-0.188	-0.218	-0.185	0.561	0.5	0.66	92.4 (0.025)
P26	-0.122	-0.123	-0.123	-0.123	-0.073	0.057	-0.011	0.579	0.68	0.88	87.4 (0.031)
P30	-0.045	-0.044	-0.044	-0.044	-0.049	-0.208	-0.150	0.647	0.63	0.82	92.2 (0.021)
P31	0.040	0.040	0.040	0.040	0.047	-0.108	0.024	0.701	0.4	0.51	89.3 (0.030)
P36	0.081	0.081	0.081	0.081	0.055	-0.172	-0.104	0.677	0.52	0.64	86.4 (0.095)
P37	0.272	0.273	0.273	0.273	0.273	0.066	0.195	0.726	0.53	0.59	79.9 (0.093)
Overall MD	0.049	0.049	0.0485	0.0485	0.051	-0.096	-0.025	0.676	0.525	0.655	87.7
(95%CI)	(-0.122, 0.149)	(-0.134, 0.149)	(-0.134, 0.149)	(-0.134, 0.149)	(-0.108, 0.164)	(-0.149, -0.019)	(0.015, 0.036)	(0.065, -0.726)	(0.479, 0.610)	(0.590, 0.771)	(86.984, 91.100)
Kruskal Wallis P value (Poor vs Good)[1]	P=0.322	P=0.322	P=0.322	P=0.322	P=0.293	P=0.155	P=0.404	P=0.841	P=0.004	P=0.003	P=0.005

* MD denotes the median of λ_j wavelet correlation[1]

Note that the threshold values for AND and RAND according to Rudge et al. (2006b) are not taken into account in classifying a patient as either a poor or good tracker. In this chapter our criterion for tracking classification is based on the patient's WCORR values at scales λ_j (j=1, 2, …, 8), their significance or otherwise (see Table 2.5), and on their "Count of NS" wavelet correlations (Table 2.5).

Figures 2.13 and 2.14 are drawn to display each patient's multivariate profile of AND, RAND and "Count of NS" (derived by 10 for axis scaling purpose), for increasing $|\lambda_1|$ for the poor trackers versus good trackers. From Figures 2.13-2.14 we observe that the profile of ("Count of NS" /10) is invariably higher for the poor trackers; and RAND, RND and $|\lambda_1|$ invariably higher for the good trackers.

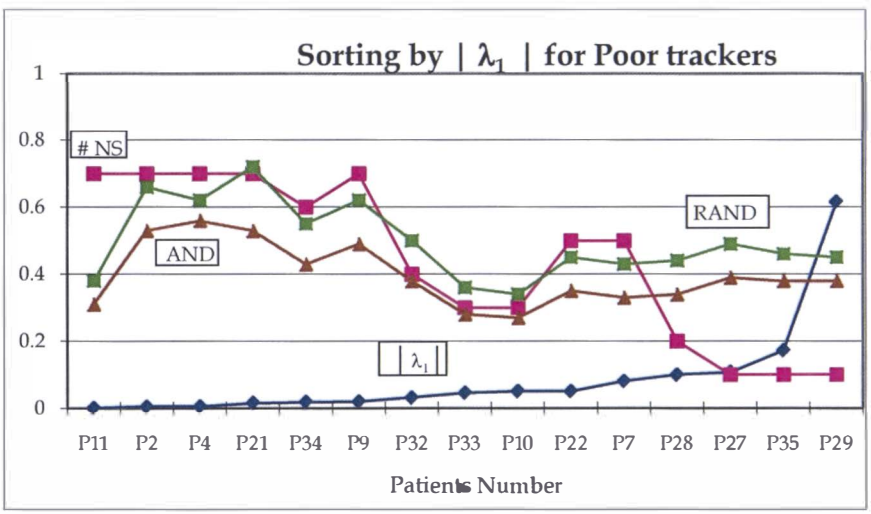


Figure 2.13 A comparison of RAND (green), AND (burgundy), the number of non-significant λ_j (divided by 10) (pink), and the modulus of λ_1 (blue) for the poor trackers (P11, P2,…, P29) sorted by increasing $|\lambda_1|$.

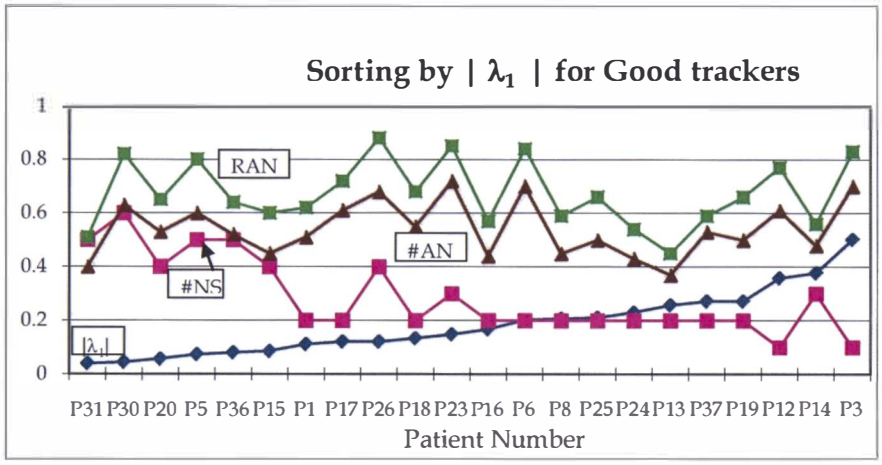


Figure 2.14 A comparison of RAND (green), AND (burgundy), the number of non- significant λ_j (divided by 10) (pink), and the modulus of λ_1 (blue) for the good trackers (P31, P30,. ., P3) sorted by increasing $|\lambda_1|$.

Table 2.7 Kruskal-Wallis tests of all wavelet and other diagnostics by wavelet based tracking group.

Variable	Poor group median	Good group median	k-w#: p value
WCORR at λ_1	-0.019	0.049	NS*
WCORR at λ_2	-0.019	0.049	NS
WCORR at λ_3	-0.019	0.049	NS
WCORR at λ_4	-0.019	0.049	NS
WCORR at λ_5	-0.008	0.051	NS
WCORR at λ_6	-0.139	-0.095	NS
WCORR at λ_7	-0. 040	-0.025	0.404+
WCORR at λ_8	0.676	0.676	NS
Absolute value WCORR at λ_1	0.046	0.159	0.001
"Count of NS"	5	2	0.001
AND	0.390	0.525	0.004
RAND	0.500	0.655	0.003
TI	85.2	87.7	0.005

k-w denotes the Kruskal-Wallis test.
* denotes non significance, $p > 0.05$
+ significant at 10% level of significance.

Table 2.7 gives the corresponding results of the Kruskal Wallis (k-w) tests for a wavelet based poor versus good tracker group effect for WCORR at scale λ_j (j=1, 2, …, 8), "Count of NS", $|\lambda_1|$, AND, RAND and TI, where grouping refers to denotes poor and good trackers. From Table 2.7 we observe that the median wavelet correlations for the first 5 wavelet scales are highly significantly lower for the poor trackers (15 of 37 patients), as are the median absolute value of the wavelet correlation at λ_1, AND, RAND and TI (p< 0.006). The median of the number of nonsignificant wavelet correlations "Count of NS" (an integer out of 8, at scales λ_j, j=1, 2,…, 8) is 5.0 for the poor trackers, significantly higher than the median of 2.0 for the good tracking group (p=0.003). It is noteworthy also that the patient specific WCORR profiles (see Figure 2.12) are good visual "signatures" of the patient's tracking status (good or poor tracking of the recorded profiles by the agitation control simulation).

Recall that 11 of our 15 wavelet DWT based poor trackers are also considered to be "poor trackers" by either or both Rudge et al's (2006b) and Chase et al's (2004) (non wavelet based) performance indicators. Kappa tests of agreement show that our DWT WCORR criterion for poor tracking developed in this chapter agrees statistically with that of Chase et al (2004) (kappa=0.2127, P=0.01) and with that of Rudge et al (kappa=0.5856, P=0,01).

Table 2.8 The signature of WCORR for the Patient 3 and Patient 4 at zero lag based WCCORR.

Patient 3 (good tracker)								
Scale	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
L. CI	-0.528	-0.539	-0.554	-0.573	-0.544	-0.398	-0.485	0.382
$\tilde{\rho}$	-0.503[++]	-0.504[+]	-0.504[+]	-0.504[+]	-0.439[+]	-0.221[+]	-0.242[+]	0.675[+]
U. CI[++]	-0.477	-0.467	-0.451	-0.428	-0.321	-0.027	0.036	0.845
Patient 4 (poor tracker)								
Scale	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
L. CI	-0.034	-0.050	-0.074	-0.107	-0.123	-0.297	-0.299	0.401
$\tilde{\rho}$	0.006	0.006	0.005	0.006	0.036	-0.077	0.023	0.724
U. CI[++]	0.045	0.061	0.085	0.117	0.194	0.151	0.339	0.886

[+] indicate WCCORR $\tilde{\rho}$ values, whose 95% CI does not include zero. These are the significant wavelet correlations for levels j (j=1, 2, … , 8).

[++] L.CI and U.CI denote the lower and upper confidence interval (CI) limits for $\tilde{\rho}$

2.3.3 Using the Wavelet Cross-Correlation (WCCORR)

We can investigate possible lead or lag relationships between a given patient's simulated versus observed (recorded) A-S profile by examining a plot of its MODWT based wavelet cross-correlation (WCCORR), according to Equation (2.8). Figure 2.15 shows this WCCORR plot for Patient 3 (P3: a good tracker) and Patient 4 (P4: a poor tracker). For Patient 3 there is large positive peak at a lag of 120 minutes for the first six wavelet scales λ_j (j=1, 2, …, 6). At scale λ_7; a large positive peak occurs at 112 minutes for Patient 3; and at scale λ_8 at a lag of 33 minutes.

We conclude that at scale λ_7 there is thus a period of 170 minutes (see Figure 2.15) for Patient 3. Likewise an examination of Figure 2.15 shows an inverse profile of peaks to troughs for Patient 4 compared to Patient 3. It is noteworthy form Figure 2.15 and that generally patients who are good trackers show common WCORR signatures or patterns, being significant at zero lag (for all scales) (see Table 2.8), not so for the poor trackers.

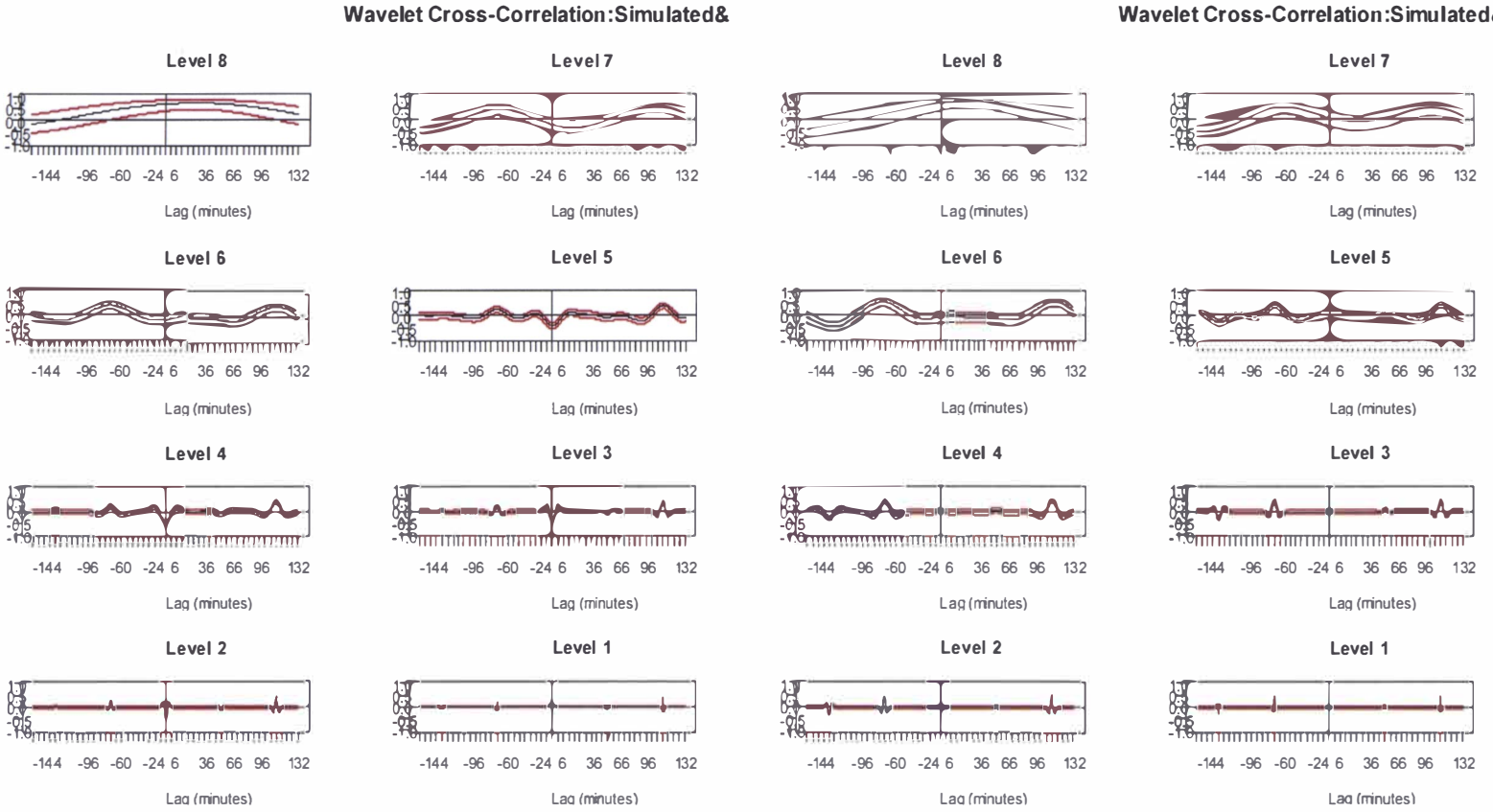


Figure 2.15 MODWT estimated wavelet Cross-Correlation between rhe simulated and recorded infusion series for lags up to ±144 minutes for Patient 3 (a good tracker) and Patient 4 (a poor tracker) with approximate 95% CI (red broken lines).

Table 2.9 Overview of results from Chapter 2

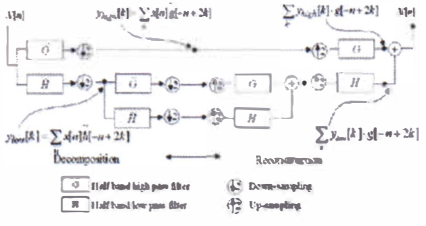
Authors	Equations or Model used	Mathematical methods used	Aims of the study and the performance indicators derived	Patient numbers with poor fit to simulated profiles (poor trackers) and the summary profile of patients with poor versus good fit
Kang et al. (Chapter 2)	See equations in the Chase et al. (2004) row below and see the schema of our approach of Chapter 2 below: **DWT Procedure (second stage)**  **DWT analysis and synthesis equations** $X=[X_1, X_2,..., X_N]$, $N=2^J$, DWT analysis equation $W=\mathcal{W}X$ $W=$ discrete wavelet coefficients, $\mathcal{W}=N \times N$ orthonormal matrix $W=\mathcal{W}X, W=\left[W_1, W_2, \cdots, W_J, V_J\right]^T$, $\mathcal{C}=\left[W_1, \mathcal{C}_2, \ldots, W_J, \mathcal{D}V_J\right]$ $X=\mathcal{W}^T W=\left[\mathcal{C}_1, W_2, \ldots \mathcal{C}_J, \mathcal{D}V_J\right]^T\left[W_1, W_2, \ldots, W_J, V_J\right]^T$ $=\sum_{j=1}^{J} W_j^T W_j + \mathcal{D}_J^T V_J \Rightarrow$ DWT synthesis equation DWT-MRA $X=\sum_{j=1}^{J} W_j^T W_j + \mathcal{D}_J^T V_J = \sum_{j=1}^{J} D_j + S_J \Rightarrow$ Additive decomposition(=MRA) $D_j=\mathcal{D}W_j^T W_j$: Portion of synthesis due to scale λ_j, jth 'detail' $S_J=\mathcal{D}_J^T V$: 'smooth' of Jth order	Discrete Wavelet Transform (DWT) (Percival & Walden, 2000) Maximal Overlap Discrete Wavelet (MODWT) (Percival & Walden, 2000) Multiresolution analysis (MRA) (Percival & Walden, 2000) DWT-MRA, MODWT-MRA	Develop wavelet correlation (WCORR) and wavelet cross-correlation (WCCORR) approach for assessing the validity of the deterministic dynamic models against the empirical agitation-sedation data of 37 ICU patients.. Provide graphical assessment tools and wavelet based numerical metrics of compatibility between the model and the data via the discrete wavelet transform (DWT), partial DWT (PDWT), maximal overlap DWT (MODWT) and via Multiresolution analysis (MRA). Investigate the lag/lead relationship between the simulated and recorded infusion series on a scale by scale basis via wavelet cross-correlation (WCCORR). Develop performance measures as follows: Modulus of the wavelet correlation at wavelet scale 1, λ_1 Count number (out of 8) of the number of non-significant wavelet correlations at scales λ_j (j=1,2,…,8). Test poor versus good fit or tracking groups via the:Kruskal Wallis test of the WCORR and indicators in 1 and 2 above for the poor versus good trackers group.	P2, P 4, P7, P9, P10, **P11, P21, P 22, P27, P28, P29, P32,P33, P34, P35** - The set of poor trackers have a median modulus of wavelet correlation at wavelet scale λ_1 of **0.046**, and 95% CI (0.017,0.093), significantly lower than the good trackers (P =0.001 k-w test)- and have a median number of non-significant wavelet correlations at scales λ_j of 5.0 (2.37, 7.0) - The set of remaining good trackers have: a significantly lower median number of non-significant wavelet correlations at scales λ_j of **2.0**, and 95% CI (2.0, 4.0)); a median modulus of wavelet correlation at wavelet scale λ_1 of **0.159**, and 95% CI (0.111, 0.232), significantly higher trackers

Table 2.8 gives the signature of $\tilde{\rho}$ and associated 95% confidence limits (L.CI, U.CI) for Patient 3 and Patient 4 based on Figure 2.15 who are deemed, to be a good and poor tracker, respectively. From the Equation 2.8, when $\tau = 0$ we obtain the MODWT estimator of the wavelet correlation between X_t and Y_t. The main results of the work in this Chapter are summarised in Table 2.9 and discussed in detail in the Conclusions (section 2.4).

2.4 Conclusions

Recently a mathematical model that captured the dynamics of a patient's A-S system was created and statistically validated for the first time by Chase et al. (2004) (see Table A.1). Further A-S models were developed by Rudge et al. (2005, 2006a, 2006b) and Lee et al. (2005).

In this study we have introduced a suite of analytic techniques based on wavelets, namely the DWT-MRA and MODWT-MRA by which to assess whether a patient's simulated agitation-sedation (A-S) status reflects their true (A-S) dynamic profile over time. Here the simulated (modeled) profiles are based on the mathematical model of A-S by Chase et al. (2004). Wavelet decomposition and reconstruction provides "signatures" of., and values for the wavelet correlations (WCORRs) and wavelet cross-correlations (WCCORRs) (at all dyadic scales), between a patient's bivariate time series, namely their simulated and their recorded A-S time series profiles.

Patients with a count of nonsignificant WCORR values greater or equal to 2 and a non significant WCORR values at scale λ_1 (level 1) are said to be a poor tracker; as are patients with a "Count of NS" less or equal to 3 and a significant negative WCORR value at scale λ_1 (level 1) and a significant negative WCORR values at scale λ_6 (level 6). See the main results of the work in this Chapter summarized in Table 2.9.

We identify 15 such poor trackers, of whom 11 are considered to be such, by either or both of the performance indicators of Rudge et al. (2006b) and of Chase et al. (2004). Our work thus shows that wavelets provide a new diagnostic tool by which to assess the agitation-sedation of ICU patients. In Chapter 3 wavelet regression and wavelet probability band type diagnostics are developed, which also mirror the poor versus good classification of this chapter's WCORR wavelets based diagnostics.

In this chapter it has also been shown that the lead or lag relationship between a patient's simulated and recorded agitation-sedation infusion series can be investigated on a scale by scale

basis via an implementation of the MODWT-MRA (using the LaDaub(8) filter). The DWT and MODWT reconstruction leads to multiresolution analysis (MRA) in which each patient's simulated or observed series can be expressed as the set of a new set of series, called details (d_j) and smooth (s). Each of the details and smooth series is associated with variations at a particular scale. The MRA is shown to point out local features of interest in both the simulated and recorded series, with the smooth MRA series offering a good visual summary of the overall long-term trends in A-S status.

Wavelet correlation and cross-correlation profiles are shown to provide "signatures" for identification of good and poor trackers. Wavelet correlation profiles of the good trackers are invariably significant at all scales (except at 32 and 64). Patients who exhibit poor tracking have WCORR profiles which are invariably non-significant at almost all wavelet scales, particularly at level 1 or scale λ

Moreover, WCCORR signatures show a common pattern for the good trackers, different to the pattern associated with poor tracking. Wavelets are also shown to visually and quantitatively discriminate patients for whom the A-S model captures their fundamental A-S dynamics, versus those, for whom this is not so. Wavelet WCORR and WCCORR modelling thus forms a possibly alternative and appropriate feedback mechanism for comparison of improved sedation administration controllers and gain.

In Chapter 3 we derive wavelet regression type diagnostics that also mirror DWT WCORR based classification, but these regressions are based on distinctly different wavelet methodologies and performance indicators.

Chapter 3

3 Density estimation and wavelet thresholding via Bayesian methods: application to agitation-sedation data

The density function is very important in statistics and data analysis. A variety of approaches to density estimation exist. Indeed the density estimation problem has a long history and many solutions (Ogden, 1997; Silverman, 1986; Walnut, 2004). A large body of existing literature on nonparametric statistics is devoted to the theory and practice of density estimation (Cline et al., 1995; Delouille et al., 2001; Ogden, 1997; Vidakovic, 1999; Walnut, 2004). The local character of wavelet functions is the basis for their inherent advantages over projection estimators – specifically that wavelets are straightforward and well localized in both space and frequency. The relevant estimation methods belong to the class of so-called projection estimators, as introduced by Čencov (1962), or their non-linear modifications. This chapter (see Sections 3.1-3.2) will trace the development of some basic methods used in density estimation. We then link these and apply wavelet methods for (density) function estimation to the ICU data of Chapter 2.

3.1 Density Estimation using Wavelet Smoothing

In order to apply wavelets to various function estimation problems in statistics, it is useful to examine some of the existing techniques in use. This provides a useful lead in to a discussion of wavelet methods for (density) function estimation, since standard techniques can be modified in a straightforward manner for use in a wavelets approach and wavelets based analysis.

3.1.1 The Haar-Based Histogram

In exploratory data analysis, it is of interest to obtain an idea of the shape of the data distribution, in the hope that interesting features (e.g. like skewness) will make themselves evident. In describing the shape of the data distribution via a histogram, we easily obtain an overall feel for the data. To illustrate the qualitative effect of bin width choice, histograms for various values of a bin width λ are plotted in Figure 3.1 for the recorded and simulated infusion profiles of Patient 12 (using the ICU data described in Chapters 1 and 2).

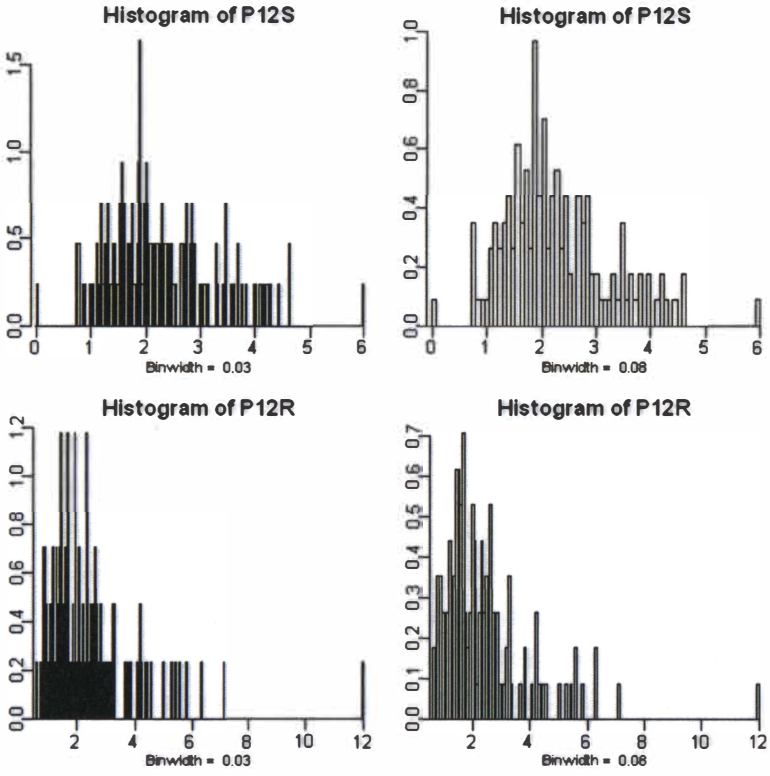


Figure 3.1 Histogram of the recorded infusion rate and simulated infusion rate for Patient 12, P12R and 12S, respectively, with varying bin widths.

A small value of bin width λ=0.03 (4-h bin width) gives a choppy looking histogram with many thin bins (Figure 3.1). Conversely, a large bin width λ=0.08 (12-h bin width) leads to over-smoothed data given the use of too few bars. When we compare the two series, they show similar features. The simulated infusion series (P12S) however is shifted a few lags to the right compared to the patient's recorded infusion series (P12R). All other patients also show similar profiles between their recorded and simulated infusion series. Most of the simulated infusion series for the 37 ICU patients have a gap in the beginning of the series, and both the recorded infusion and the simulated infusion series are not symmetric.

We now discuss specialised versions of histograms that can be constructed using the Haar wavelet basis. Important theoretical properties of this estimator are discussed by Engel (1990). The Haar wavelet approach and histogram leads naturally to density estimators with smoother wavelet bases and lend themselves to histogram estimators, as we require. Given the Haar scaling function, on the left hand side of Equation (3.1), and then applying the usual dilation and translation gives,

$$\phi(t)=\begin{Bmatrix}1, \text{ if } 0\le t<1\\ 0, \text{otherwise}\end{Bmatrix}\Rightarrow \phi_{j,k}(t)=\begin{Bmatrix}2^{j/2}, 2^{-j}k\le t<2^{-j}(k+1)\\ 0, \quad \text{otherwise}\end{Bmatrix}. \tag{3.1}$$

We can then count up the number of data points that lie within a particular interval, say, $\left[2^{-j}k, 2^{-j}(k+1)\right)$ using the quantity $2^{-j/2}\sum_{i=1}^{n}\phi_{j,k}(t_i)$.

Now for any $t\in R$ and $j\in Z$,

$$2^{-j}\left[2^{j}t\right]\le t<2^{-j}\left(2^{j}t+1\right), \tag{3.2}$$

where $[t]$ denotes the greatest integer function of t. Therefore the number of data points that lie in the same interval as any real number t can be computed by

$$2^{-j}\sum_{i=1}^{n}\phi_{j,\left[2^{j}t\right]}(t_i)=\sum_{i=1}^{n}\phi\left(2^{j}t_i-2^{j}t\right). \tag{3.3}$$

The histogram density estimator with origin 0 and bins of width 2^{-J} is given by

$$\tilde{f}_J(t)=\frac{1}{n}2^{J/2}\sum_{i=1}^{n}\phi_{J,\left[2^{J}t\right]}(t_i). \tag{3.4}$$

This estimator can be regarded as being the best estimator of the density f on the approximation space V_J, where V_J is defined by Equation (2.4) in Chapter 2. Construction of histograms using the Haar basis, then leads to more general wavelet density estimators. The decomposition algorithm can be applied to Equation (3.1) and the resultant histogram can be written in terms of the Haar wavelets as follows:

$$\tilde{f}_J(x)=\sum_{k}\tilde{c}_{j_0,k}\Phi_{j_0,k}(x)+\sum_{j=j_0+1}^{J-1}\sum_{k}\tilde{d}_{j,k}\Psi_{j,k}(x). \tag{3.5}$$

The Haar-based histograms are given in Figure 3.2 for level 1, 2, 3, and 4, using the simulated infusion rate (red colour) and the recorded infusion rate (black colour) for Patient 12 and Patient 18, who are good trackers according to the DWT wavelet criterion in Chapter 2. Figure 3.2 shows a similar distribution between the patient's recorded and simulated infusion rates - skewed right for both patients (P12 and P18). Correspondingly Figure 3.3 presents the simulated data and recorded A-S data of Patient 2 and Patient 27 who are judged to be poor trackers according to the

wavelet diagnostics in Chapter 2. Each patient's simulated and recorded series are clearly of a differing distribution type to each other (Figure 3.3) for these so-called poor trackers. Figures 3.2 and 3.3 are clearly more informative in comparison to the histogram given in Figure 3.1 where the former density estimates are based on the Haar wavelet basis (Figure 3.2-3.3). These graphical comparisons allow us to visualise the differences in distribution between the poor and good tracking patients.

3.1.2 Estimation by Smooth Wavelets

Estimating density functions using smooth wavelets can be performed in the same way for any orthogonal series. A natural application of wavelets, this estimation procedure results from a straightforward extension of the Haar-based histogram approach (Ogden, 1997). The same approach that was used to estimate a density in terms of the Haar basis above can also be used with smooth wavelet bases, as we now illustrate.

Let ϕ and ψ be an orthogonal scaling function and mother wavelet pair that generate a series of approximating spaces $\{V_j\}_{j\in\mathbb{Z}}$. Then $f(x)$, which is a square integrable density function, is

$$f(x)=\sum_k c_{j_0,k}\phi_{j_0,k}(x)+\sum_{j>j_0}\sum_k d_{j,k}\psi_{j,k}(x), \tag{3.6}$$

where j_0 represents a coarse level of approximation. Haar coefficients are estimated using

$$\hat{c}_{j,k}=\langle \tilde{f},\phi_{j,k}\rangle=\frac{1}{n}\sum_{i=1}^{n}\phi_{j,k}(X_i) \tag{3.7}$$

$$\hat{d}_{j,k}=\langle \tilde{f},\psi_{j,k}\rangle=\frac{1}{n}\sum_{i=1}^{n}\psi_{j,k}(X_i). \tag{3.8}$$

From Equation (3.7) and (3.8) given above, the wavelet estimator for f at level $J\geq j_0$ is given by

$$\begin{aligned}\hat{f}_J(x)&=\sum_k \hat{c}_{j_0,k}\phi_{j_0,k}(x)+\sum_{j>j_0}\sum_k \hat{d}_{j,k}\psi_{j,k}(x)\\&=\sum_k \hat{c}_{J,k}\phi_{J,k}(x).\end{aligned} \tag{3.9}$$

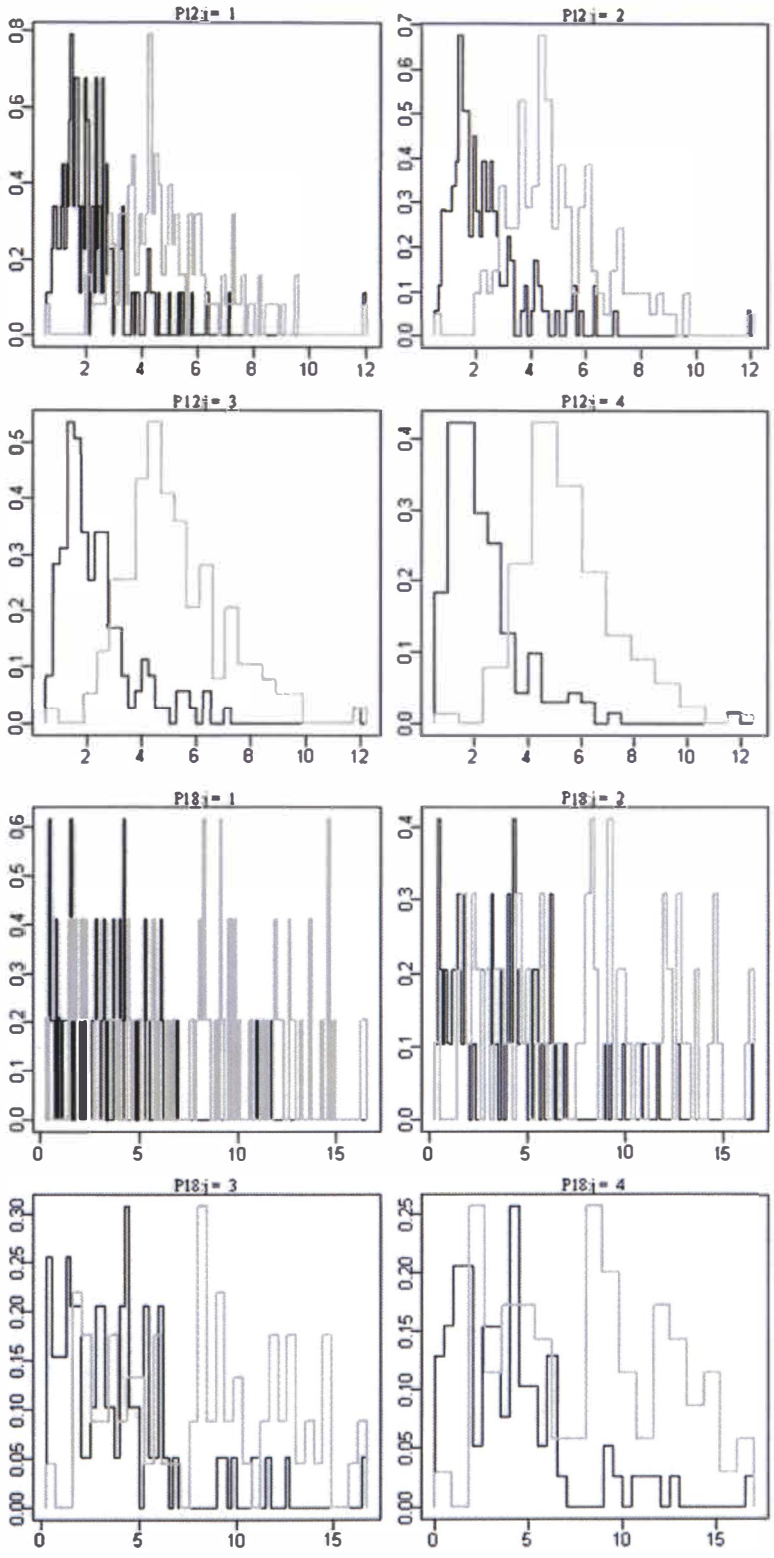


Figure 3.2 Haar-based histogram of the simulated series (light) and recorded A-S series (dark) for varying resolution levels for two "good trackers": Patients 12 (top 4 plots) and 18 (bottom 4 plots).

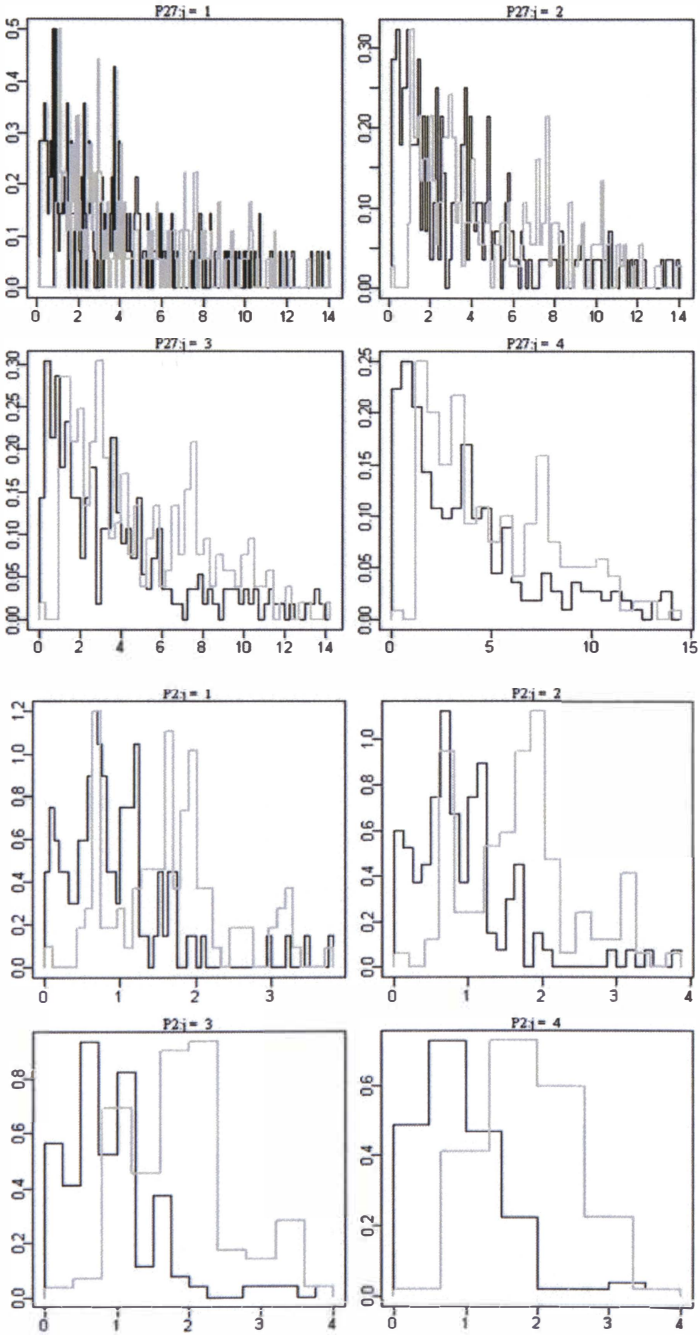


Figure 3.3 Haar-based histogram of the simulated series (light) and recorded series (dark) for varying resolution levels for two "poor trackers": Patients 27 (top 4 plots) and 2 (bottom 4 plots).

The smoothing parameter in Equation (3.9) is the index J of the highest level considered. Smooth wavelet-based density estimates are plotted in Figure 3.4 for level 4, 6, and 8, using the Patient 4's simulated infusion and recorded infusion rate data (via Daub (4) which denotes the Daubechies wavelet filter of length 4). We sampled 2048 ($=2^{11}$) data points without loss of any generality from the original data of Patient 4. Figure 3.5 shows Patient 29's smooth wavelet-based density estimates (using Daub (4)). Figures 3.4 and 3.5 indicate that Patent 4 (P4) is potentially a poor tracker and Patient 29 (P29) a good tracker, since the original and wavelet smooth densities are similar in P29, but not for P4. This gives support for the two (poor versus good) groupings in Table 2.6. Note that when level J is increased, abrupt jumps disappear but thereby we also over-smooth and lose information needed to distinguish between the R and S series. In Figure 3.5, Patient 29 shows similar features between its recorded and simulated data so Patient 29 is possibly a good tracker. We qualify and quantify good and poor tracking in Section 3.3 via wavelet density estimation methods.

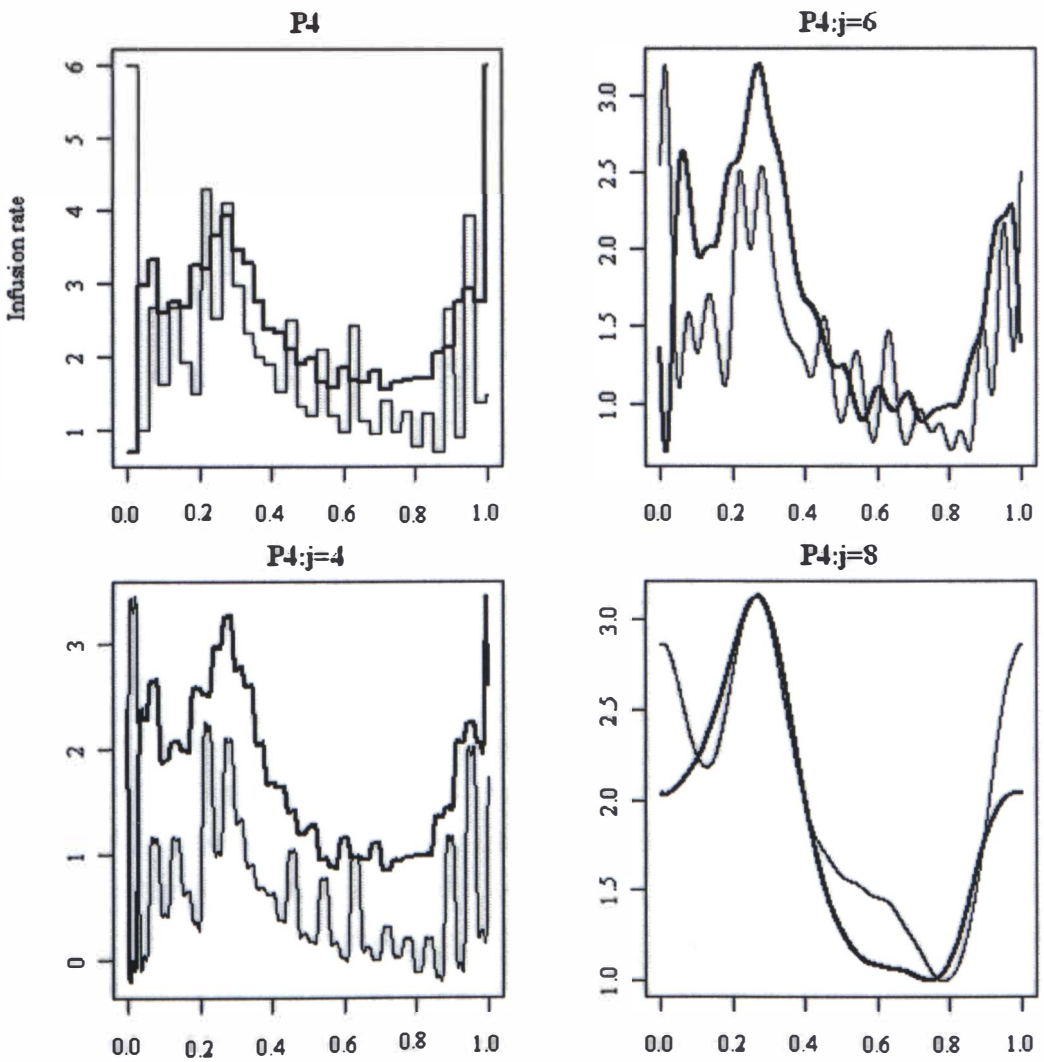


Figure 3.4 Smooth wavelet-based density estimates for P4's recorded data (light) and simulated data (dark) using the Daubechies wavelet (Daub4) with sub- sample N=2048 and for different choices of J.

3.2 Density Estimation using Wavelet Shrinkage

3.2.1 Wavelet Shrinkage (threshold)

To quantify the relationship between the two variables $(x_i,\ y_i)$, we can employ the standard regression model as follows,

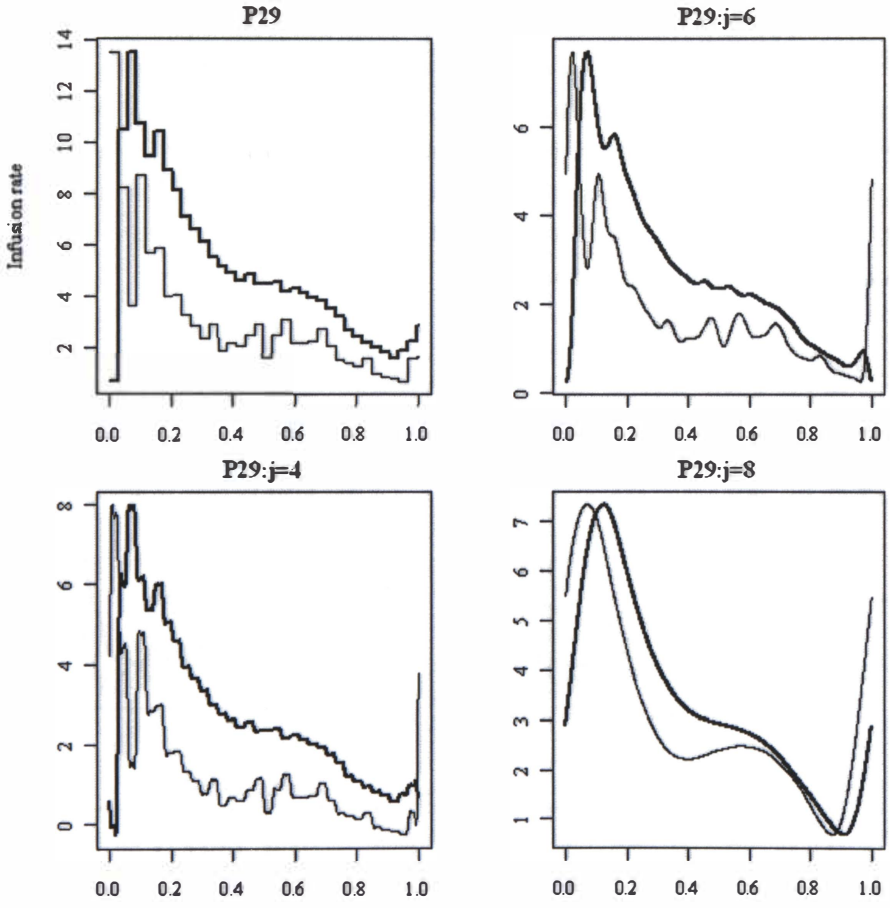


Figure 3.5 Smooth wavelet-based density estimates for P29's recorded () and simulated data (dark) using the Daubechies wavelet (Daub4) with N=2048 and for different choices of *J*.

$$y_i = f(x_i) + \varepsilon_i,\ i = 1, \ldots, n, \tag{3.10}$$

where the ε_i's are independent and identically distributed $\mathrm{N}(0,\ \sigma^2)$ random variables. It will be assumed that the design points $(x_1, \ldots, x_n)$ are equally spaced, and, without further loss of generality, that they lie on the unit interval: $x_i = 1/n,\ i = 1, \ldots, n$. In Section 3.1.2 our approach

constitutes projecting the raw estimator f onto the approximating space V_J, for any choice of the smoothing parameter J, which represents a linear estimation approach. In contrast to this, Donoho & Johnstone (1994), offer a non-linear wavelet based approach to nonparametric regression. Their approach begins with computing the DWT of the data y_i, by generating a new data set of empirical wavelet coefficients, with which to represent the underlying regression function f. The Donoho and Johnstone (1994) estimation procedure has three main steps, as follows.

First, transform the data y_i to the wavelet domain by applying a DWT. If d is the DWT of f, and $d' = \left[c'_{0,0}, d'_{0,0}, \ldots, d'_{J-1,2^J-1} \right]^T$ the vector of empirical coefficients, then we have a sequence of wavelet coefficients $d' = \mathrm{d} + \varepsilon'$, where ε' is a vector of n independent $\mathrm{N}\left(0, \sigma^2\right)$. In the second step, the true coefficients d are estimated by applying the thresholding rule to the empirical coefficients d' to obtain estimates $\tilde{d}$. Finally, the sampled function values f are estimated by applying the inverse DWT (IDWT) to obtain $\tilde{f} = W^T \tilde{d}$, where W^T is the transpose of an orthonormal $n \times n$ matrix. We can then represent this DWT as the sum

$$\tilde{f}(x) = \tilde{c}_{0,0}\phi(x) + \sum_{j=0}^{J} \sum_{k=0}^{2^j-1} \tilde{d}_{j,k}\psi_{j,k}(x). \qquad (3.11)$$

This procedure is schematised in Figure 3.6.

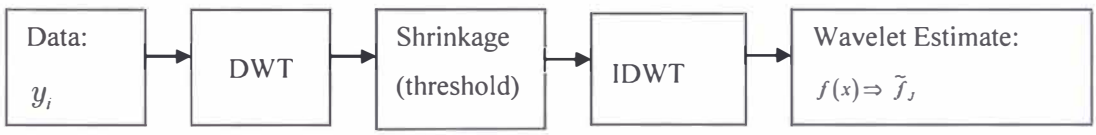


Figure 3.6 Wavelet shrinkage (threshold) procedure.

A technique for selective wavelet reconstruction similar to this general approach was proposed by Weaver et al. (1991) to remove random noise from magnetic resonance images. However, Donoho and Johnstone (1994) developed the technique from the statistical point of view by considering selective wavelet reconstruction as a problem in multivariate normal decision theory. The largest true coefficients should thus be included in a selective reconstruction. Indeed in estimating an unknown function it is natural to include only coefficients larger than some specific threshold value. A large coefficient is taken to mean one that is large in absolute value (Donoho & Johnstone, 1994). Choosing a very large threshold will make it difficult for a coefficient to be

judged significant and included in the reconstruction, finally resulting in over smoothing. On the other hand, a very small threshold value will allow many coefficients to be included in the reconstruction, and so result in under-smoothed estimates. Donoho and Johnstone (1994) propose two methods of global wavelet thresholding, namely a universal threshold where $\lambda = \sqrt{2\log n}$, and the minimax threshold. The minimax threshold values (λ) given in Table 3.1 are reproduced from Donoho and Johnstone's paper (1994).

Table 3.1 Minimax thresholds (λ) for various sample sizes, from Donoho & Johnstone (1994)

n	λ	n	λ
64	1.474	2048	2.414
128	1.669	4096	2.594
256	1.860	8192	2.773
512	2.047	16384	2.952
1024	2.232	32768	3.131

Figure 2.4 in Chapter 2 shows that the simulated infusion profile tracks the average recorded infusion profiles, rather than following the instantaneous variations in the infusion rate. The severe local variations in the recorded infusion rate are likely to be a result of the variability and subjectivity inherent in a "human-based" feedback protocol, and are not observed in the simulations, where the simulated nurse is 100% consistent. Therefore, to objectively compare the patient's recorded and simulated infusion profiles, the recorded profile is first smoothed using wavelet thresholding (see Section 3.2.2). Although linear methods are available for use in wavelet thresholding estimation, we prefer to concentrate on non-linear methods because of their ability to adjust to rapidly non-stationary features in an observed process (Donoho & Johnson, 1998).

From Figures 2.5 and 2.6 in Chapter 2, the DWT is implemented prior to performing wavelet thresholding on all 37 patients. The assumption of a dyadic length of the time series is not always available to us in practice. A natural approach would then be to precondition the original data set, so as to get a set of values of length 2^J for some positive integer J. The resulting preconditioned data is then plugged directly into any standard DWT. In Table A.3 of Appendix A, we see that most of the 37 patient data is not to the power of two. One obvious remedy is to pad the series with values and increase its length to the next power of two. There are several choices for the value of these padded coefficients. In this study, we pad with zeros to increase the size of the data set to the next larger power of two or some other highly composite number, and then apply the DWT. Using the minimax estimator with soft thresholding to the simulated infusion profile of Patient 2 we obtain the profile in Figure 3.7.

3.2.1.1 Numerical approach: Wavelet Time Coverage Index (WTCI)

The most commonly used criterion to obtain a successful wavelet estimator of the signal $\tilde{x}$ in estimating x is the mean square error (MSE) (Donoho, 1995). In this study, we devise a variant based on the development of the smoothed recorded infusion. This then lays the foundation for our development of a Wavelet Time Coverage Index (WTCI), which is a quantitative parameter indicating how well the simulated infusion profile represents the patient's average recorded infusion profile over the entire time series. This uses wavelet coefficients on a scale by scale basis. The WTCI is defined as follows:

$$\text{WTCI} = \left\{1 - \frac{\sum_{j,k} \left| \tilde{d}_{j,k} - d_{j,k} \right|}{\sum_{j,k} \tilde{d}_{j,k}}\right\} \times 100 \qquad (3.21)$$

where $\tilde{d}_{j,k}$ is given Equation (3.11) and $d_{j,k}$ is the DWT of f in Equation (3.10).

A WTCI =100% therefore represents perfect tracking. This scenario arises when the simulated infusion (mathematically modelled) profile is identical to that of the wavelet smoothed recorded infusion profile.

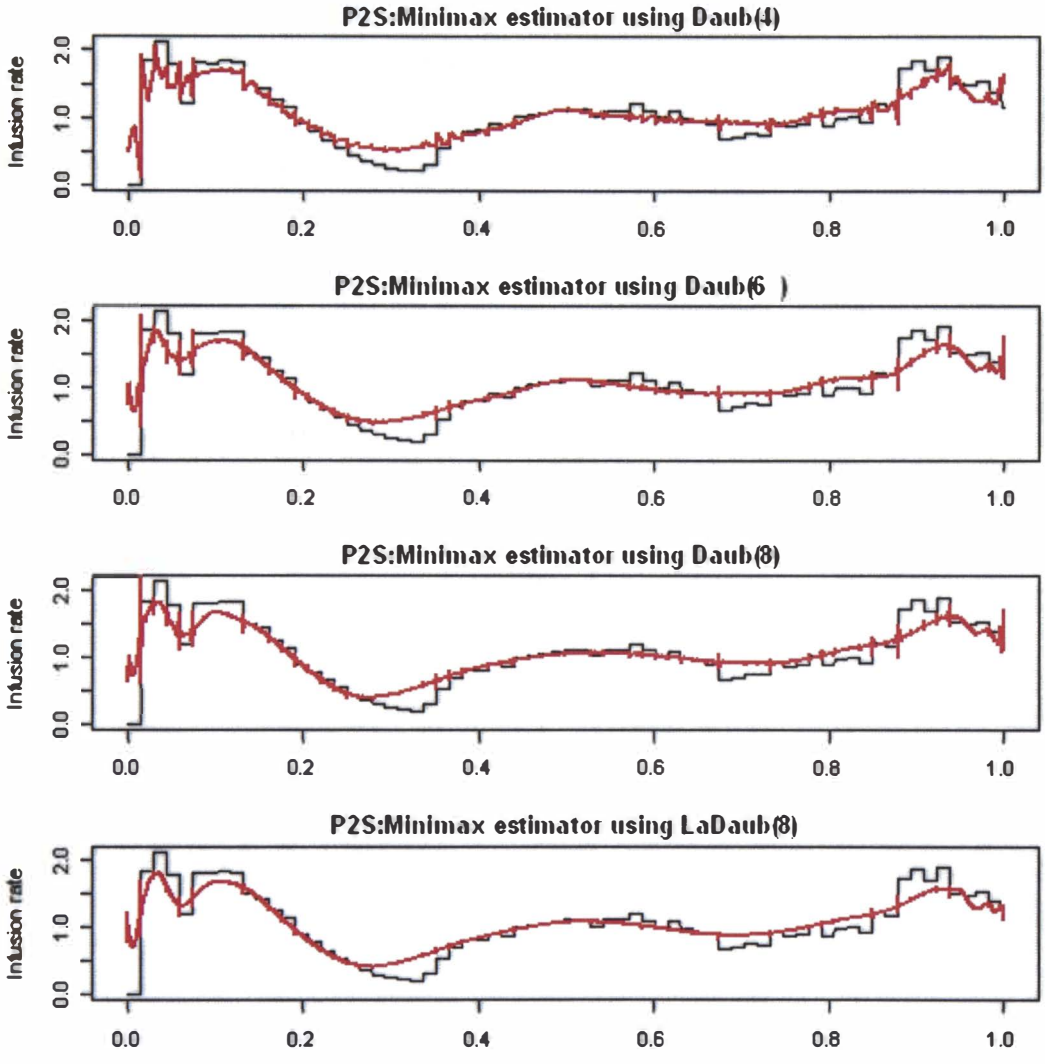


Figure 3.7 Minimax estimator applied to Patient 2's simulated profile (the thick line represents the wavelet threshold estimator of the simulated infusion rate and the thin line that of the recorded infusion data). The soft thresholding rule was used to obtain all estimates.

3.2.1.2 Numerical approach: Average Normalized Wavelet Density (ANWD) and the Relative Average Normalized Wavelet Density (RANWD)

In this section we propose so-called wavelet analogues of the AND and RAND diagnostics of Rudge et al. (2006a; 2006b). Let $y_t = \{y_1, y_2, \ldots, y_n\}$ be the output data produced by a proposed model. These are often called the simulated data. Define the average normalized wavelet density (ANWD) of y_t as

$$\mathrm{ANWD}(y_t) = \frac{1}{n}\sum_{t=1}^{n} \frac{\tilde{f}_t(y_t)}{\max\left(\tilde{f}_t\right)}, \tag{3.22}$$

where $\max\left(\tilde{f}_t\right)$ denotes the maximum value of the wavelet density function $\tilde{f}_t$, which is estimated by wavelet smoothing via Equation (3.9) at time t. Thus, ANWD is an average of normalized densities, where each component in the sum is the value of $\tilde{f}_t$ at y_t, normalized by $\max\left(\tilde{f}_t\right)$). At time t, the normalized wavelet density equals 1 when y_t coincides with the point where $\tilde{f}_t$ is maximum. An infusion profile that coincides with the maximum wavelet density at every time point would therefore have ANWD equal to 1. Whereas the value of ANWD of an infusion profile distant from the high-density regions would approach 0.

Finally, ANWD (y) is calibrated using the ANWD of the wavelet smoothed recorded infusion data, denoted by $\tilde{y}$, giving the relative average normalized wavelet density (RANWD):

$$\mathrm{RANWD} = \frac{\mathrm{ANWD}(X)}{\mathrm{ANWD}(\tilde{y})}. \tag{3.23}$$

RANWD indicates the value of ANWD (y) relative to a typical realisation in the form of $\tilde{y}$, from the density profile. Therefore, the RANWD statistic estimates how probabilistically alike the model outputs are to the smoothed data, and hence the degree of compatibility between the model (simulated) and the empirical (recorded) data. For example, a RANWD of 0.6 may be interpreted as the scenario where the model outputs are 60% similar, on average, to the wavelet smoothed data. Greater similarity means greater compatibility between the simulated and the empirical values for a given patient.

3.2.2 Wavelet thresholding via Bayesian methods

Recall the estimation of a regression equation (Equation (3.10)) in Section 3.2.1 for an observed data vector $y_1, y_2, \ldots, y_n$ satisfying

$$y_i = f(x_i) + \varepsilon_i, \qquad i = 1, \ldots, n,$$

where the ε_i 's are independent and identically distributed $\mathrm{N}(0, \sigma^2)$ random variables, assuming that $(x_1, \ldots, x_n)$ are fixed points. The formulation of f in terms of a wavelet basis is characteristically sparse, and distills most of the signal (in the data) into a small number of large

valued coefficients, whereas the noise is distributed "evenly" across the coefficients given the orthonormality requirement of the wavelet basis. The data is denoised by a preselected thresholding rule so as to reject "small" coefficients and keep large coefficients considered to contain the signal.

A method to approximate the posterior distribution of each $f(x_i)$, via the same prior as utilised by the BayesThresh method (of Abramovich et al., 1998 and Nason. 2008) is now considered. Posterior probability intervals of any nominal coverage probability can be computed correspondingly.

For Haar wavelets, the scaling function and mother wavelet are $\phi(t) = I(0 \le t < 1)$ and $\psi(t) = I(0 \le t < 1/2) - I(1/2 \le t < 1)$, respectively, where $I(\cdot)$ is the indicator function. The square of the Haar wavelet is just the Haar scaling function, then $\psi^3(t) = \psi(t)$ and $\psi^2(t) = \psi^4(t) = \phi(t)$. Furthermore, we can have, and $\psi^4_{j,k}(t) = 2^{3j/2}\phi_{j,k}(t)$. All these terms can be incorporated into the pyramid algorithm which includes scaling function coefficients. We represent $\psi^r_{j,k}(t)$ by scaling functions and Haar wavelets, and also are able to use a comparable method for alternative wavelet bases. Following the development in Herrick (2000), approximate a general wavelet $\psi^r_{j_0,0}, (0 \le j_0 \le J-m)$, by

$$\psi^r_{j_0,0} \approx \sum_l e_{j_0-m,l}\phi_{j_0-m,l}(t) \tag{3.24}$$

for r = 2, 3, 4, where m is a positive integer. The choice of m follows below. Since scaling functions (instead of wavelets), as the span of the set of scaling functions at a given level j, are the same as that of the union of $\phi(t)$ and the wavelets at levels 0, 1,..., j-1. Furthermore, if scaling functions $\phi_{j,k}(t)$ are used to estimate some function $Q(t)$, and both $\phi_{j,k}(t)$ and $Q(t)$ have at least ν derivatives, then the mean squared error in the approximation is constrained by the upper limit $C2^{-\nu j}$, where C is some positive constant, (see, for example, Vidakovic (1999) and Abramovich et al. (1998)).

To approximate $\psi^r_{j_0,k}(t)$ for some fixed j_0, we simply compute $\psi^r_{j_0,0}(t)$ using the pyramid algorithm (Daubechies, 1992), then take the DWT and set the coefficients $e_{m_0,l}$ to be equal to the

scaling function coefficients $e_{m_0,k}$ at level m_0, where $m_0 = j_0 + m$. Recall that the wavelets at level j are simply shifts of each other; from $\psi_{j,k}(t) = \psi_{j,0}\left(t - 2^{-j}k\right)$, hence

$$\psi^r_{j_0,0} \approx \sum_l e_{m_0,l-2^m k}\phi_{m_0,l}(t). \tag{3.25}$$

As we are assuming periodic boundary conditions, the $e_{m_0,l}$ can be cycled periodically.

Given the localised nature of wavelets, the coefficients $e_{m_0+1,l}$ employed to approximate $\psi^r_{j_0-1,0}(t)$ can be found by inserting 2^{m_0} zeros into the vector of $e_{m_0,l}$

$$e_{m+1,l} = \begin{cases} \sqrt{2}e_{m_0,l} & l = 0,\ldots 2^{m_0-1}-1 \\ 0 & l = 2^{m_0-1},\ldots,2^{m_0+1}-2^{m_0-1}-1 \\ \sqrt{2}e_{m_0,l-2^{m_0}} & l = 2^{m_0+1}-2^{m_0-1},\ldots,2^{m_0-1}-1 \end{cases} \tag{3.26}$$

The approximation Equation (3.24) cannot be implied for wavelets at the m finest levels $J-m,\ldots,J-1$. These wavelets are written in terms of both scaling functions and wavelets at the finest level of detail, level J-1, via the block-shifting method as delineated above.

From Equation (3.11) we have that $[f_i \mid y]$ is the convolution of the posteriors of the wavelet coefficients and the scaling coefficient given by,

$$[f \mid y] = \left[c_{0,0} \mid c''_{0,0}\right]\phi(t_i) + \sum_j \sum_k \left[d_{j,k} \mid d''_{j,k}\right] \psi_{j,k}(t_i). \tag{3.27}$$

If X and Y are independent random variables and a and b are real constants, then

$$\kappa_r(aX+b) = \begin{cases} a\kappa_1(X)+b, & r=1 \\ a^r\kappa_r(X), & r=2,3,\cdots \end{cases} \tag{3.28}$$

and we have by the additivity property

$$\kappa_r(X+Y) = \kappa_r(X)+\kappa_r(Y), \quad r \in \mathbb{Z}. \tag{3.29}$$

for all r. More details on cumulants can be found in Barndorff-Nielsen & Cox (1989) or Stuart & Ord (1994).

Applying Equations (3.28) and (3.29) to Equation (3.27) gives $\left[f_i \mid y\right]$ (where $f_i = f\left(x_i\right)$) can be estimated from its cumulants as

$$\kappa_r\left(f_i \mid y\right) = \kappa_r\left(c_{0,0} \mid c''_{0,0}\right)\phi^r_{j,k}\left(t_i\right) + \sum_j \sum_k \kappa_r\left(d_{j,k} \mid d''_{j,k}\right)\psi^r_{j,k}\left(t_i\right). \qquad (3.30)$$

The first cumulant, $\kappa_1\left(y\right)$, is the mean of y, the second cumulant, $\kappa_2\left(y\right)$, is the variance of y, $\kappa_3\left(y\right)/\kappa_2^{3/2}\left(y\right)$ is the skewness, and $\kappa_4\left(y\right)/\kappa_2^2\left(y\right)+3$ is the kurtosis and that the third cumulant $\kappa_3\left(y\right)$ and the fourth cumulant $\kappa_4\left(y\right)$ are zero if y is Gaussian.

From Equation (3.25) and (3.30), we can now re-write the fourth cumulant as follows,

$$\begin{aligned}\kappa_r\left(f_i \mid y\right) &= \kappa_r\left(c_{0,0} \mid c^*_{0,0}\right)\phi^r_{j,k}\left(t_i\right) + \sum_j \sum_k \kappa_r\left(d_{j,k} \mid d^*_{j,k}\right)\psi^r_{j,k}\left(t_i\right) \\ &= \kappa_r\left(c_{0,0} \mid c^*_{0,0}\right)\phi^r_{j,k}\left(t_i\right) + \sum_{j,k}\left\{\kappa_r\left(d_{j,k} \mid d^*_{j,k}\right)\sum_l e_{j-3,l}\phi_{j-3,l}\left(t_i\right)\right\} \\ &= \sum_{j,k}\rho_{j,k}\phi_{j,k}\left(t_i\right) \qquad (3.31)\end{aligned}$$

for $\kappa_r\left(y\right)$ the r^{th} cumulant of y, and for suitable coefficients, $\rho_{j,k}$, acquired via the IDWT algorithm which incorporates scaling function coefficients so as to assess this sum (Barber et al, 2002; Abramovich et al.,1998; Nason, 2008).

Bayesian wavelet regression estimates have developed, including priors on the wavelet coefficients $d_{j,k}$, which are updated by the observed coefficients $d''_{j,k}$ to obtain posterior distributions $\left[d_{j,k} \mid d''_{j,k}\right]$ (see Equation (3.27)). $\tilde{d}_{j,k}$ (point estimates) can be computed from these posterior distributions and the Inverse Discrete Wavelet Transform (IDWT) used to estimate $f\left(x_i\right)$.

The Bayesian wavelet shrinkage rules have used mixture distributions as priors on the coefficients, to model the notion that a small proportion of coefficients contain substantial signal. Chipman et al. (1997) and Crouse et al. (1998) considered mixtures of two normal distributions, while Abramovich et al. (1998), Clyde et al. (1998) and Clyde & George (2000) used mixtures of a normal and a point mass.

The BayesThresh method of Abramovich et al. (1998) puts the independent priors on the coefficients:

$$d_{j,k} \sim \gamma_j N\left(0, \tau_j^2\right) + \left(1-\gamma_j\right)\delta(0),\ j = 0,\ldots,J-1;\ k = 0,\ldots,2^j - 1,$$

where $0 \le \gamma_j \le$, $\delta(0)$ is a point mass at zero and $d_{j,k}$ are independent. The hyperparameters are assumed to be of the form $\tau_j^2 = C_1 2^{-\alpha j}$, $\gamma_j = \min\left(1, C_2 2^{-\beta j}\right)$ for non-negative constants C_1 and C_2 chosen empirically from the data and α and β selected by the user. Abramovich et al. (1998) show that choices α and β of correspond to choosing priors within certain Besov spaces, incorporating prior knowledge about the smoothness of $f(x_t)$ in the prior. See more details in Abramovich et al. (1998) and Chipman & Wolfson (1999).

3.3 Application to the ICU agitation-sedation data

3.3.1 Choice of Wavelet filter and Bootstrap: WTCI

In order to judge the reliability of the wavelet time coverage index (WTCI) for a given patient's infusion profile, the moving blocks bootstrap was utilized (Efron & Tibshirani, 1993). A total of 1000 bootstrap realizations were generated for each of the patient's recorded infusion profiles. A wavelet time coverage index (WTCI), as defined in Equation (3.21), can then be evaluated for each bootstrap realization, providing a collection of 1000 values of the WTCI. The median WTCI and its standard error, SE, can then be reported for each patient (Hettmansperger & McKean, 1998) (see Table 3.2).

When the DWT is implemented via a pyramid algorithm (Mallat, 1989), we need to choose the appropriate wavelet filter (basis). The choice of a wavelet basis function is important when analysing a given time series for two reasons. First, the length of a DWT determines how well it approximates an ideal band-pass filter, which in turn dictates how well the filter is able to isolate features to specific frequency intervals. Secondly, as illustrated in the MODWT MRAs shown in Figures 2.7 and 2.8 in Chapter 2, the wavelet basis function is being used to represent information contained in the time series of interest and should imitate its underlying features. A reasonable

overall strategy is thus to use the shortest width of wavelet filters $L = 4, 8$ and longer wavelet filters $L = 10, 12$, as both choices give reasonable results in this ICU application.

Table 3.2 presents the data size (time series length) and median WTCI and its standard error for each of the 37 patients involved in this study. The poor trackers as classified in Chapter 2 are bolded in the first column. From Table 3.2, we note the following. Some poor trackers (Patients 9, 11, 22, 27, 32, 34) have high values of the median WTCI and some good trackers (Patients 1, 14) have a low WTCI. The reason for this is padding, although a reasonable solution to produce data of the size power of two, is the fact that it dilutes the signal near the end of the original data set, since the filters are not applied evenly. This means that multiplying by a signal element, constrained to have magnitude zero, is equivalent to omitting the filter coefficient. Then the orthogonality of the transform is not strictly maintained. To overcome this problem we change the current *minutes driven length* of the data set to hourly, then we apply Bayesian wavelet shrinkage methods using a universal threshold (see Section 3.2.2).

Table 3.2 Wavelet Time Coverage Index summary for the 37 patients.

Patient no.	Data size[1] (Min)	Median WTCI	SE
1	3601	78.55	0.538
2[2]	6421	87.14	0.294
3	6541	87.85	0.106
4	4921	87.94	0.076
5	2941	88.89	0.103
6	5701	88.73	0.104
7	3901	84.78	0.324
8	10561	93.61	0.037
9	8581	93.28	0.046
10	20701	88.46	0.053
11	6721	92.46	0.085
12	8521	91.15	0.323
13	5161	91.37	0.091
14	3001	82.09	0.449
15	4981	92.09	0.072
16	13621	94.57	0.073
17	5941	90.27	0.086
18	4681	93.83	0.036
19	7921	96.34	0.012
20	9661	90.49	0.088
21	3721	83.07	0.685
22	9661	91.85	0.056
23	3481	85.07	0.300
24	8461	92.41	0.058
25	3841	93.44	0.082
26	3901	85.49	0.275
27	13441	93.66	0.039
28	12241	89.47	0.051
29	3241	89.3	0.262
30	3661	85.81	0.092
31	18301	94.34	0.022
32	15181	95.82	0.020
33	25261	95.63	0.036
34	8101	93.77	0.070
35	12721	87.64	0.018
36	3481	92.17	0.059
37	7501	90.79	0.079
Max		96.340	0.685
75th%		93.525	0.184
Med		90.790	0.079
95%CI		(88.746,92.386)	(0.058,0.092)
25th%		87.745	0.049
Min		78.550	0.012
SE		0.85	0.032

[1] Data size indicates the length of the time series sampled.

[2] Bolded Patient no. indicates a poor tracker by the DWT, WCORR and WCCORR diagnostics developed in Chapter 2.

3.3.2 Alternative WTCI measure via Bayesian Wavelet Thresholding

In Section 3.3.1, we usethe minimax estimator (Donoho & Johnson, 1998) to obtain patient specific WTCI measures. In Section 3.3.2 we calculate the WTCI using Bayesian wavelet thresholding (Clyde et al., 1998; Vidakovic, 1998; Chipman et al., 1997; Abramovich et al., 1998; Nason, 2008). Table 3.3 gives all values of the results of the WTCI measure obtained by using Bayesian wavelet thresholding as computed by the R and WaveThresh software packages (Nason, 2008) (using the LaDaub (8) filter). The relative total dose, defined as the total drug dose delivered by the simulation, as a percentage of the total recorded drug dose, is also presented for each patient in Table 3.3.

Figure 3.8 presents the box and whisker plot for values of the WTCI per patient from the bootstrap realizations (per patient). Each box and whisker in Figure 3.8 displays two main components of information. First, the median represents a measure of how well the agitation-sedation simulation models the recorded infusion profile on average. Second, the spread of the box and whisker plot provides an indication of how reliable that particular WTCI median is (Figure 3.8). Therefore, a box and whisker plot displaying a high median and low spread is indicative of good tracking performance (for a given patient), indicating that the simulation captures their essential A-S dynamics well. Plots displaying high spread indicate poor WTCI reliability, and no conclusion regarding the performance of the simulation can be made. Poor reliability may be caused by insufficient data, and also by the choice of wavelet filters, the chosen thresholding method, or the model simulation.

Table 3.3 Wavelet Time Coverage Index (WTCI) summary for the 37 patients.

Patient no.	Median WTCI	SE	Relative total Dose (%)
1	74.0	0.378	87.8
2[+]	77.4	0.461	85.9
3	76.6	0.264	86.1
4	79.2	0.650	86.6
5	79.3	0.324	90.3
6	84.6	0.697	87.5
7	75.5	0.083	81
8	84.3	0.381	92.1
9	88.1	0.359	89.1
10	74.9	0.175	90.2
11	72.0	0.395	87.8
12	83.8	0.261	86.5
13	82.8	0.166	90.2
14	76.6	0.248	88.5
15	86.3	0.145	90.7
16	84.8	0.431	90.4
17	83.3	0.430	85.2
18	84.2	0.161	95.0
19	86.5	0.341	91.0
20	80.5	0.303	90.4
21	72.1	0.498	87.8
22	77.5	0.171	89.5
23	79.8	0.673	91.7
24	85.1	0.127	89.9
25	88.9	0.075	91.2
26	82.8	0.061	88.5
27	75.7	0.663	87.5
28	75.0	0.202	90.6
29	73.4	0.442	77.0
30	83.3	0.224	94.5
31	82.6	0.146	90.0
32	82.6	0.271	91.2
33	78.5	0.430	90.4
34	85.9	0.193	89.0
35	74.1	0.361	88.1
36	79.1	0.293	81.6
37	78.6	0.283	88.9
Max	88.9	1.065	95.0
75^{th} %[▽]	84.3	0.43	90.4
Median	79.8	0.293	89.1
95%CI: Median	(77.57, 83.23)		
25^{th} %[▽]	76.14	0.173	87.5
Min	71.9	0.061	77.0
SE	0.9	0.033	0.6

[+] Bold patient no. indicates a poor tracker as defined in Chapter 2.
[▽] 75^{th}% and 25^{th} % denote the relevant percentiles.

Table 3.3 and Figure 3.8 show that for most patients the variability of the bootstrap realizations is very small, indicating high reliability of the median WTCI. Some poor trackers (Patients 9 and 34) have high values of WTCI with small spread. Some good trackers (Patients 1, 3, 23) have low values of WTCI with large spread. Therefore no conclusion can be drawn from the WTCI values of patients 1, 3 and 23. The observed high values of the median WTCI for the infusion profiles certainly indicate the validity of the developed model. The median patient specific WTCI across all patients of 79.8% with a 95% interpolated confidence interval (CI) of (77.57%, 83.23%) and with a range [71.9%, 88.9%], while dependent upon the definition of WTCI indicates significant merit of the developed mathematical model of Chase et al. (2004) and its physiological validity. Note that the median and 95% confidence interval for WTCI for the poor trackers (see later for the wavelets based criterion for poor versus good tracking) is 76.56% (74.89%, 81.69%) and for good trackers is 82.79% (79.13%, 84.10%) (Kruskal Wallis p=0.041) (see Table 3.7 for further details).

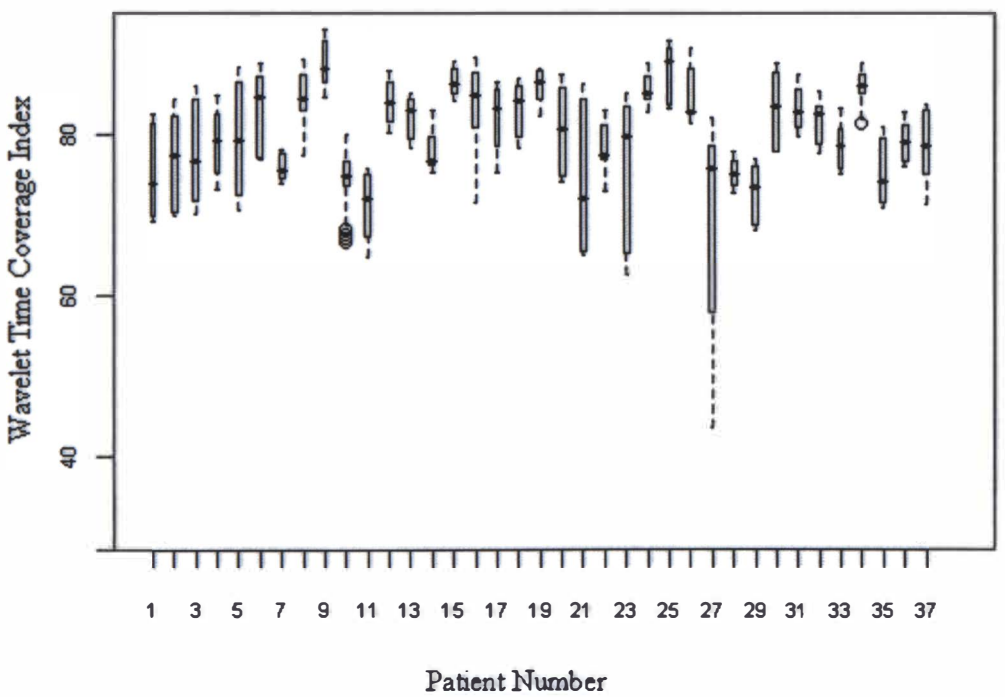


Figure 3.8 Box and whisker plot of WTCI index for the 37 patients.

Furthermore, the overall patients median relative total dose is 89.1 % (Table 3.3) with a range [77.0%, 95.0%] indicating that the simulated and recorded total drug doses are similar, with the simulator consistently administering slightly less than 100% of the recorded actual sedative dose. Such slightly decreased levels are linked with the sudden-response nature of the recorded infusion profiles, in contrast to the consistent, smooth quality of the simulated infusion profile (Chase et al.,

2004; Rudge et al., 2005; 2006a; 2006b). These features are chiefly a result of the consistency of the computer-implemented simulation in contrast to the inherent variation between different nurses' assessment of patient agitation and appropriate feedback of sedation (Chase et al., 2004; Rudge et al., 2005; 2006a; 2006b).

Numerical metrics comparing the model outputs to the recorded data are now developed using the Bayesian densities determined from the recorded data. By using Bayesian wavelet shrinkage methods we can create the wavelet probability band (WPB). A 90% value for the WPB implies that for at least 90% of the time, the estimated mean value of the recorded infusion rate for a given patient lies within the band. A 90% wavelet probability band is constructed for each of the 37 patient profiles, and the time and duration of any deviations from the wavelet probability band are recorded. Figure 3.9 shows the 90% WPB for good trackers (Patients 8, 23) and for poor trackers (Patients 9, 34). The circle represents the hourly recorded infusion rate, the thin line represents the 90% WPB and the solid thick line represents the simulated profile (Figure 3.9). The brief spikes which occur in the bands are typical of wavelet regression methods. These can be smoothed out by using different α and β, but this risks over-smoothing the data because we lose some information. From Abramovich (1998), setting $\alpha=0.5$, $\beta=1$ seems to be the best practical approach in Bayesian smoothing methods. Therefore, we used $\alpha=0.5$ and $\beta=1$, in this study and also utilized Daubeches' least asymmetric wavelet with eight vanishing moments, namely LaDaub (8), as this a widely used wavelet and is applicable to a wide variety of data types.

Table 3.4 presents the time that the simulated infusion rate lies within the 90% wavelet probability band (denoted by WPB 90%). Similar types of trends are observed for all 37 ICU patients, as demonstrated by the generally high values of WPB 90%, in the second column of Table 3.4. With the exception of twelve (12) patients (Patients 2, 4, 7, 9, 10, 21, 22, 27, 28, 32, 33, 34), all simulated infusion profiles lie within the wavelet probability band at least 71% of the time. Of the 12 previously mentioned patients with WPB 90% $\geq$ 71% the following all twelve (12) patients (Patients 2, 4, 7, 9, 10, 21, 22, 27, 28, 32, 33, 34) were also deemed to be "poor trackers" according to the WCORR diagnostics developed in Chapter 2 (see Table 2.5 and later Table 3.6). The main reason for the reduced total time within the WPB for the 12 poor trackers is a consistently poor performance throughout the total length of the simulation. This feature is observed in Figure 3.9 for P9 and for P34.

Table 3.4 Simulated infusion profile compatibility: wavelet probability band for the 37 patients

Patient no.	WPB 90%	ANWD	RANWD
1	95.31	0.537	0.553
2	64.06	0.431	0.499
3	96.88	0.632	0.737
4	59.38	0.338	0.475
5	93.75	0.495	0.504
6	95.31	0.659	0.980
7	67.19	0.417	0.455
8	87.50	0.567	0.688
9	57.81	0.343	0.412
10	66.80	0.300	0.388
11	77.34	0.423	0.434
12	84.38	0.622	0.662
13	73.44	0.442	0.504
14	96.88	0.449	0.476
15	89.06	0.702	0.761
16	82.81	0.596	0.770
17	85.94	0.506	0.566
18	93.75	0.548	0.558
19	74.22	0.759	0.780
20	96.88	0.487	0.581
21	65.62	0.407	0.413
22	65.62	0.422	0.455
23	92.19	0.288	0.341
24	71.10	0.655	0.635
25	89.06	0.635	0.670
26	96.88	0.600	0.601
27	47.27	0.368	0.608
28	50.78	0.501	0.540
29	82.81	0.343	0.394
30	96.88	0.554	0.597
31	87.50	0.562	0.669
32	68.36	0.326	0.362
33	58.79	0.373	0.500
34	48.44	0.505	0.551
35	96.10	0.371	0.533
36	75.00	0.573	0.763
37	79.69	0.448	0.607
Min	47.27	0.288	0.341
Median	82.81	0.495	0.552
Max	96.88	0.759	0.981

Note: A bold patient no. indicates a poor tracker

Table 3.4 also shows that most patients for whom the simulated infusion profiles lie in the WPB less than 90% of the time also have large maximum departure time differences between their recorded and simulated infusion rates. This indicates that the simulated infusion profile deviates from the recorded infusion profile over some particular period, and takes some time before tending towards the recorded infusion rate again (Figure 3.9, Patients 9 and 34).

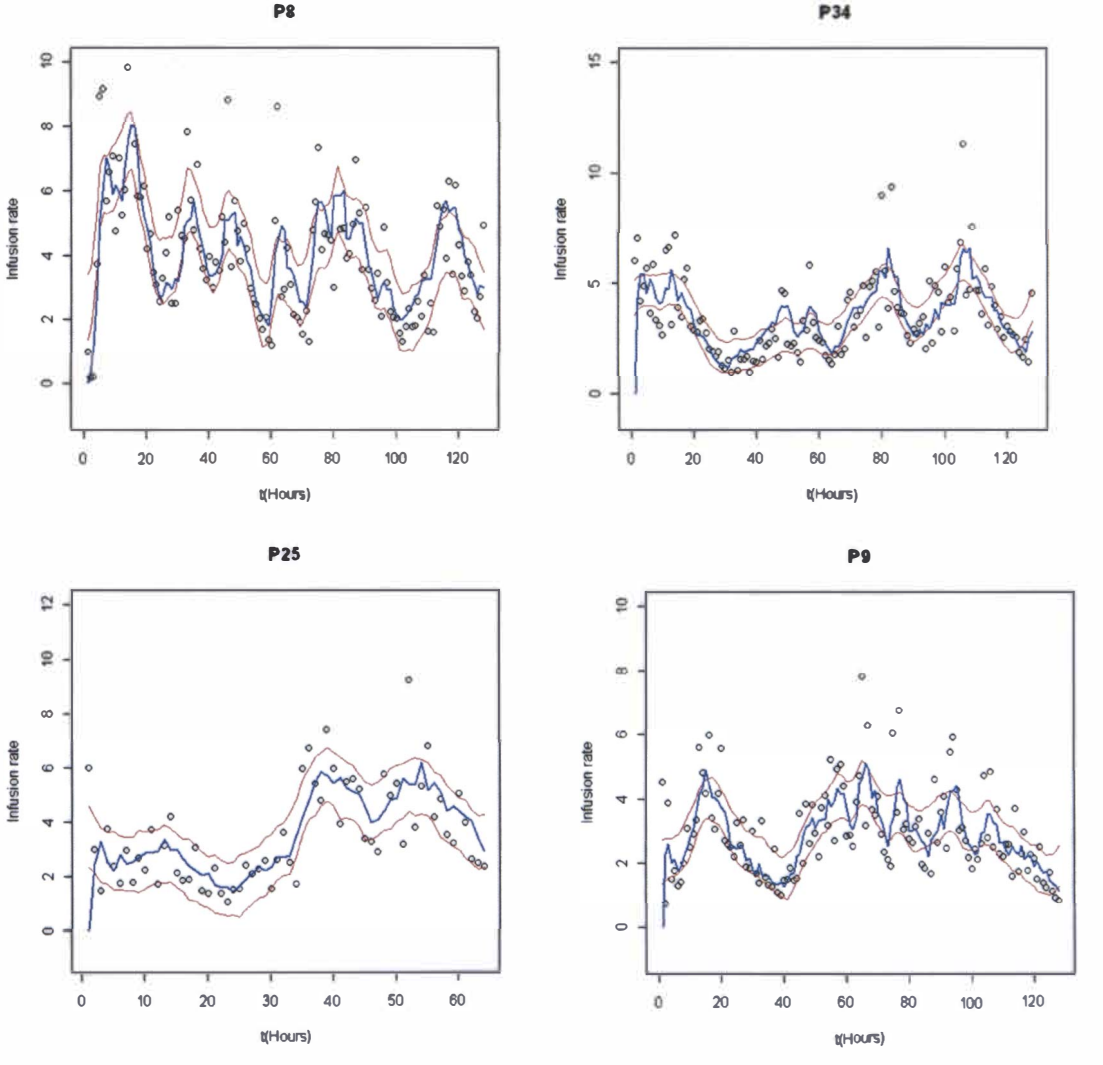


Figure 3.9 90%Wavelet Probability Bands (WPB's) (thin lines) with simulated infusion profile (thick line) for Patients 8, 25 (P8, P25: good trackers, LHS) and Patients 9, 34 (P9, P34: poor trackers, RHS).

Patients 25 and 34 can now be used to illustrate some of the concepts linking and differentiating the different wavelet measures in Tables 3.3 and 3.4. For example, Patient 25 has the maximum WTCI of 88.9% with a high value of relative total dose of 91.2% (Table 3.3), and exhibits low spread in the bootstrapped realization in Figure 3.8. Table 3.4 shows that Patient 25's simulated infusion rate lies 89.06% of the time within the 90% WPB, showing good performance or

tracking, as evident in Figure 3.9 . By contrast, the simulation infusion profile for Patient 34 has a WTCI of 85.9% and relative total dose of 89.0% (Table 3.3), and exhibits low spread in the bootstrapped realization but interestingly has a low WPB 90% value of only 48.44% (Figures 3.8-3.9 and Tables 3.3 and 3.4). Hence whilst the WTCI values of Patient 25 and Patient 34 are both high, 88.9% and 85.9%, respectively; it is only the WPB 90% measure, not the WTCI measure, that distinguishes between Patient 25 (WPB 90% = 89.06%) and Patient 34 (WPB 90% = 48.44%).

Recall from Chapter 2 that Patient 25 is deemed to be a good tracker and Patient 34 a poor tracker. Therefore, while the simulation infusion rate is outside the band for 51.56% of the time for Patient 34, a large maximum departure time between Patient 34's recorded and simulated infusion rate occurs in that period (see Table 2.1 in Section 2.3.3); even though the extent to which Patient 34's profile leaves the band is small, indicated by its relatively high WTCI value of 85.9% (Table 3.3).

3.3.3 ANWD and RANWD measures

The higher the percentage of full recorded time profile that the simulated infusion profile lies within the wavelet probability band, the better the model captures the specific dynamics of the agitation-sedation system for that patient. While our WPB approach (Figure 3.9) is graphically useful, it however is deficient in that it is not an objective numerical measure of how close the simulated infusion profile is to the empirical data. The percentage WPB (WPB 90%) does not serve this purpose of objective quantification, because it simply quantifies visual proximity, by means of artificial hard boundaries, and ignores the fact that the in-band region does not have the same probabilistic significance everywhere.

New numerical metrics, namely ANWD and RANWD, comparing the model outputs to the recorded data are therefore now developed using the posterior densities determined from the smoothed recorded data. The density profile is more informative than the wavelet probability band (WPB) since the latter does not discriminate between regions of high or low probability within a band. The density profile is used in this section, as mentioned earlier, to compute the numerical measures of ANWD and RANWD, which were defined in Section 3.2.1.2, so that objective comparisons of model performance can be made across different patients.

The ANWD values for the simulated infusion rates is the average of these normalized density values over all time points for a given patient. Similarly, the RANWD values for the smoothed infusion rate is obtained by superimposing the smoothed values by using the first cumulant from a normal posterior distribution onto the same density profile, after which RANWD can be readily computed (see Section 3.2.1.2).

Columns 3 and 4 in Table 3.4 present our new performance measures of ANWD and RANWD (defined by Equations (3.22) and (3.23), respectively) per patient. Recall that RANWD measures how probabilistically similar the model outputs are to the smoothed observed data, and hence is a measure of the degree of compatibility between the simulated and the empirical data. It should be noted that as the mathematical model is deterministic, its outputs do not belong to the same probabilistic mechanism that generated the data, hence RANWD is an extremely stringent measure. Consequently, consistently high RANWD values that are close to 1.00 are not expected, even for a good simulation model. A reasonable and practical threshold for adequate model performance is RANWD $\geq$ 0.5, which infers that the model outputs are more alike than not, to the smoothed data. Justification for our 0.5 threshold for RANWD is given below.

The posterior distribution of the regression curve is used as the density profiles for all patient simulations for the original model of Chase et al. (2004). An overall median RANWD of 0.552 with range [0.341, 0.981] supports the visual finding of closeness with a statistically based objective measure (see Figure 3.9). Only 13 patients have a RANWD$\leq$ 0.5, indicating that 68% of the simulated infusion profiles were, on average, a close representation of the smoothed recorded infusion profile. More specifically 9 patients have RANWD values from 0.412-0.499, and 4 have RANWD values between 0.341-0.394. These 13 patients are (P2, P4, P7, P9, P10, P11, P14, P21, P22, P23, P29, P32, P33) all of whom, except P14, P23 and P29, were also identified as poor trackers by the WCORR and WCCORR DWT diagnostics in Chapter 2. Of the three afore-mentioned patients and including P29 (i.e. P14, P23, P29), we note that P29 and P33 were judged to be poor trackers according to Rudge et al. (2006b).

Table 3.5 Comparison between the WPB, ANWD and RANWD values and the TIB, AND, and RAND measures from Rudge's Physiological Model (Rudge et al., 2006b)

Patient no.	WPB model			Rudge's Physiological Model		
	WPB 90%	ANWD	RANWD	TIB 90%	AND	RAND
1	95.31	0.537	0.553	96	0.51	0.62
2	64.06	0.431	0.499	90	0.53	0.66
3	96.88	0.632	0.737	97	0.70	0.83
4	59.38	0.338	0.475	93	0.56	0.62
5	93.75	0.495	0.504	97	0.60	0.80
6	95.31	0.659	0.980	95	0.70	0.84
7	67.19	0.417	0.455	67	0.33	0.43
8	87.50	0.567	0.688	90	0.45	0.59
9	57.81	0.343	0.412	89	0.49	0.62
10	66.80	0.300	0.388	53	0.27	0.34
11	77.34	0.423	0.434	59	0.31	0.38
12	84.38	0.622	0.662	96	0.61	0.77
13	73.44	0.442	0.504	85	0.37	0.45
14	96.88	0.449	0.476	95	0.48	0.56
15	89.06	0.702	0.761	95	0.45	0.60
16	82.81	0.596	0.770	91	0.44	0.57
17	85.94	0.506	0.566	91	0.61	0.72
18	93.75	0.548	0.558	92	0.55	0.68
19	74.22	0.759	0.780	90	0.50	0.66
20	96.88	0.487	0.581	91	0.53	0.65
21	65.62	0.407	0.413	95	0.53	0.72
22	65.62	0.422	0.455	83	0.35	0.45
23	92.19	0.288	0.341	95	0.72	0.85
24	71.10	0.655	0.635	91	0.43	0.54
25	89.06	0.635	0.670	86	0.50	0.66
26	96.88	0.600	0.601	92	0.68	0.88
27	47.27	0.368	0.608	84	0.39	0.49
28	50.78	0.501	0.540	76	0.34	0.44
29	82.81	0.343	0.394	90	0.38	0.45
30	96.88	0.554	0.597	97	0.63	0.82
31	87.50	0.562	0.669	74	0.40	0.51
32	68.36	0.326	0.362	74	0.38	0.50
33	58.79	0.373	0.500	67	0.28	0.36
34	48.44	0.505	0.551	84	0.43	0.55
35	96.10	0.371	0.533	70	0.38	0.46
36	75.00	0.573	0.763	83	0.52	0.64
37	79.69	0.448	0.607	92	0.53	0.59
Min	47.27	0.288	0.341	53	0.27	0.34
Median	82.81	0.495	0.552	90	0.49	0.60
Max	96.88	0.759	0.981	97	0.72	0.88

Note: A bold patient no. indicates a poor tracker by the wavelet based criteria in Chapter 2.

Table 3.6 Patient numbers of poor trackers according to study criterion

Kang WPB diagnostics (Chapter 3)	Kang WCORR diagnostics (Chapter 2)	Chase et al. (2004) diagnostics	Rudge el al. (2006b) diagnostics
-	-	-	-
2	2	-	-
-	-	-	-
4	4	-	-
-	-	-	-
-	-	6	-
7	7	7	7
-	-	-	-
9	9	9	-
10	10	-	10
11	11	-	11
-	12	12	-
-	-	-	13
14	-	-	-
-	-	-	-
-	-	-	-
-	-	17	-
-	-	-	-
-	-	-	-
-	-	-	-
-	21	21	-
22	22	-	22
23	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-
27	27	27	27
28	28	-	28
-	-	-	29
-	-	-	-
-	-	-	-
32	32	-	-
33	33	-	33
-	34	34	-
-	35	-	35
-	-	-	-
-	-	-	-
Total: N_1=13	Total: N_2=15	Total: N_3=8	Total: N_4=10

Indeed Figures 2.11 and 2.12 provide support for our wavelet criteria (see Table A.1 in the appendix). Note that specifically Patients 14 and 23 (good trackers) have very low RANWD values (see Table 3.4) but have high WPB 90% and WTCI values (Tables 3.3-3.4). Their WPB and density profiles revealed specific regions where the model for these patients did not appear to capture the observed patient dynamics. In some cases, this occurs when the absence of a stimulus or low drug concentrations coincide with an agitation level that was decreasing (but not close to zero), causing the patient's agitation to remain at a constant non-zero level, despite their recorded infusion rate dropping to near zero.

Rudge's Physiological Model (Rudge et al., 2006b) also found that 27 out of 37 patients (73%) have values of RAND ≥ 0.5 (see Tables 3.5 and 3.6), 5 patients have $0.43 < \text{RAND} < 0.49$, and 3 have $0.34 < \text{RAND} < 0.38$. Lee's (2005) model likewise found that 27 of 37 patients (73%) have RAND values greater than 0.57 have RAND values ranging from 0.43 to 0.49, and 3 patients have RAND values between 0.34 and 0.38 (see Tables 3.5 and 3.6). The performance metrics of AND and RAND are defined in Table A.2 and their patient specific values are also given in Table 3.5.

Table 3.7 Summary of wavelet based performance statistics: poor versus good trackers.

	WPB %	WTCI	ANWD	RANWD
Poor trackers				
Min	47.27	71.95	0.03	0.36
Max	77.34	88.05	0.50	0.61
Range	30.07	16.01	0.21	0.25
Mean	61.56	77.97	0.39	0.47
95% CI of Mean	(65.79, 67.32)	(74.72, 81.22)	(0.36, 0.44)	(0.42, 0.51)
Median	64.84	76.56	0.41	0.46
95% CI of Median	(52.63, 67.09)	(74.89, 81.69)	(0.34, 0.43)	(0.41, 0.53)
Good trackers				
Min	58.79	73.37	0.29	0.34
Max	96.88	88.94	0.76	0.98
Range	38.09	15.57	0.47	0.64
Mean	85.31	81.36	0.53	0.60
95% CI of Mean	(80.85, 89.77)	(76.62, 83.09)	(0.48, 0.57)	(0.55, 0.66)
Median	87.50	82.79	0.55	0.59
95% CI of Median	(82.72, 93.79)	(79.13, 84.09)	(0.45, 0.60)	(0.53, 0.67)
Kruskal-Wallis test P value	0.001	0.041	0.001	0.001

Tables 3.5 and 3.6 allow comparison between Rudge's (2006b) values of AND, and RAND with our WPB model diagnostics (WPB, ANDWD, RANWD). According to Table 3.5 and Table 3.6, six(6) patients (namely, patients 7, 10, 11, 22, 27, 28) are poor trackers and 20 patients are good trackers as ascertained from using both the WPB diagnostic and also Rudge's Physiological Model (2006b) criterion.

Table 3.7 gives summary statistics of the wavelet density based metrics for the poor versus good trackers as defined in Section 3.4 (using the threshold criterion WPB 90% $\leq$ 0.50). The poor trackers have significantly lower median values of WPB 90% (64.84% versus 87.50%) ($p \leq$ 0.002), WTCI (76% versus 82.79%) ($p \leq$ 0.041), ANWD (0.41 versus 0.55) ($p \leq$ 0.002) and RANWD (0.46 versus 0.59) ($p \leq$ 0.002) compared to the good trackers.

3.4 Discussion

In this chapter a nonparametric approach for statistical assessment of the validity of deterministic dynamic models of empirical agitation-sedation data is developed, based on wavelet smooth regression and its density estimation. Our methods yield visual graphical assessment tools as well as numerical metrics of compatibility between the modelled and observed A-S data.

Generally a Bayesian approach has been suggested for assessing a parametric statistical model by constructing a wavelet probability band for the proposed model and then checking whether the nonparametric regression curve lies within the band. However, for nonparametric *wavelet* regression there is no wavelet probability band that can be constructed. Our contribution in this chapter is to construct a wavelet *probability band* for the nonparametric wavelet regression curve and check whether the proposed model lies within the band. This so-called wavelet probability band (WPB) also gives a useful visual tool to assess the compatibility of the mathematically simulated and recorded A-S time series profiles per patient.

Moreover, the density profile is used to define and compute two numerical measures, namely the average normalized wavelet density (ANWD) and the relative average normalized wavelet density (RANWD) - measures of closeness between the recorded and simulated infusion rate. These graphical and quantitative approaches are useful in both identifying and distinguishing between the good and poor trackers. Our approach is also a valuable tool for detecting regions where the simulated infusion rate (modelled A-S) performs poorly, thus providing ways to improve and distil mathematical models of agitation-sedation.

In Chapter 3, we developed a density estimation approach via wavelet smoothing for assessing the validity of deterministic dynamic models (giving the simulated profiles) against the empirical (recorded) data. This wavelet density approach provides graphical assessment and numerical metrics of compatibility between the model and the recorded agitation-sedation data. The WTCI and 90% WPB give strong support for the model in Chapter 2 and vice versa.

Wavelet modelling in Chapter 2 and Chapter 3 demonstrate that the models of the recent A-S studies of Chase et al., (2004), Rudge et al., (2005; 2006a; 2000b) and of Lee et al., (2005), are suitable for developing more advanced optimal infusion controllers (see Table A.1 and Table 2.9 and Table 3.8). Through the wavelet modelling, we can depict the dynamics between A-S profiles. These offer significant clinical potential of improved agitation management and reduced length of stay in critical care. Specifically Rudge et al. (2006b) derived a physiological model which provided a platform to develop and test future semi- automated sedation management controllers.

Overall, our work (see Table 3.8, Table 2.9 and Table A.1) shows that it is possible to create and evaluate wavelet models of the agitation-sedation system and use them to develop wavelet diagnostics by which to evaluate agitation management, where the latter represents a trade-off between the benefits of low patient agitation versus the cost of high infusion rates and increased total dose requirements (Rudge et al., 2003; 2005; 2006a; 2006b). This study clearly shows that wavelet modelling can capture the fundamental dynamics of the agitation-sedation system. Wavelets are thus suitable for clinical implementation in ICU agitation and sedation control.

Table 3.8 Overview of Agitation-Sedation studies of ICU patients

Authors	Equations or Model used	Mathematica methods usec	Aims of the study and the performai indicators derived	Patient numbers w poor fit to simula profiles (p trackers) and summary profile patients with p versus good fit
Kang et (Chapter 3	See the schema of the approach developed in Chapte below Wavelet shrinkage (threshold) procedure Data X_t → DWT → Threshold (minimax) → IDWT → $f(x) \Rightarrow \tilde{f}_J$ (wavelet estimate) $f(x) = \sum_k c_{j_0,k}\phi_{j_0,k}(x) + \sum_{j>j_0}\sum_k d_{j,k}\psi_{j,k}(x)$: a square intergrable density functi ϕ : orthogonal scaling function, ψ : mother wavelet $\hat{f}_J(x) = \sum_k \hat{c}_{j_0,k}\phi_{j_0,k}(x) + \sum_{j>j_0}\sum_k \hat{d}_{j,k}\psi_{j,k}(x)$: wavelet estimator for $f(x)$ at J $\hat{c}_{j,k} = \frac{1}{n}\sum_{t=1}^{n}\phi_{j,k}(X_t)$, $\hat{d}_{j,k} = \frac{1}{n}\sum_{t=t}^{n}\psi_{j,k}(X_t)$	Density estimation (Ogden, 199 Wavelet thresholding BayesThresh methods (Abramovich al., 19 1998b) Wavelet shrinkage (threshold) (Gencay et 1. 2002; Ogd 1997)	Develop a density estimation approach wavelet smoothing for assessing validity of deterministic dynamic mod (simulated profiles) against the empiric recorded data. Construct a wavelet probability ba (WPB). Provide graphical assessment numerical metrics of compatibi between the model and the recor agitation-sedation data. Develop performance measures as follo 1. Average normalized wavelet den (ANWD). 2. Relative average normalized wav density (RANWD). 3. Median of the Wavelet Time Cover Index (WTCI).	**P2, P4, P7, P9, P P11, P14, P22, P P27, P28, P32, P33** Of the subset of patients above, 9 h RANWD values fr (0.412-0.499), and have even lo RANWD val between (0.341-0.3 The remaining patients all h RANDWD val greater than 0.5.

Chapter 4

4 Wavelet Denoising of Images

4.1 Introduction

Many images are contaminated with noise, either because of their acquisition process, or because of naturally occurring phenomena. One reason for this is that many images are acquired under less than ideal conditions. Several important processing operations, such as contrast enhancement, histogram equalization, and edge enhancement, work much better if random noise is absent. These contamination problems have been recognised for a long time and as yet there is no optimal solution for removing noise. A trade-off between the removed noise and the blurring in the image always exists.

Denoising algorithms are used primarily to minimize the noise level, while conserving the inherent features of the image. Conventional denoising methods are founded on the technique of smoothing and stem from the assumption that for a majority of images, the signal's energy can be represented by a small number of transform coefficients, whilst the noise contributes to the high frequency and traditionally non-significant coefficients. The standard denoising method for Fourier based techniques, is the Wiener filter (Wiener, 1964; Gencay et al., 2002), which attenuates the part of the spectrum that is of high frequency, but subsequently also eliminates several image details as well. On the other hand, suppressing coefficients in the domain of the wavelet transform has much potential, in that the localized nature of wavelet coefficients makes them more appropriate for application to locally adaptive methods for image processing (Chang, 2000a, 200b).

Relatively recently there has been much study on filtering and wavelet coefficients thresholding, due to the ability of wavelets to afford a suitable basis for separating the noisy versus the image signal (see Daubechies, 1992; Donoho, 1995; Donoho & Johnson, 1995a, 1995b, 1994). Thresholding of the Discrete Wavelet Transform (DWT) coefficients is the cornerstone of wavelet-based methods, with an extensive number of image processing algorithms being based on the DWT approach (Gencay et al., 2002; Percival & Walden, 2000). The transform coefficients within the subbands can be locally modelled as independent identically distributed random variables with a Generalized Gaussian Distribution (GDD) (Mallat, 1989). In that sense, the denoised coefficients may be evaluated by a Minimum Mean Square Error (MMSE) estimator, in

terms of the noise coefficients and the variances of signal noise. The denoised coefficients are statistically estimated in small regions for every subband instead of applying a global threshold (Mihcak et al., 1999). In Chang et al. (2000b) a similar spatially adaptive model for wavelet image coefficients was used to perform image denoising via wavelet thresholding. Since the study of Donoho and Johnstone (1995a; 1994), there has been much work on defining the threshold levels and their type (see the literature review in Chapter 1).

An image denoising algorithm attempts to remove the noise from the image. Ideally, the resulting denoised image will not contain any noise or added artefacts. Denoising of natural images using wavelet techniques, is very effective because of their ability to capture the energy of a signal in a few energy transform values (Chang, 2000a, 200b).

The methodology of general wavelet denoising of the discrete wavelet transform based image involves the following three steps:

1. Transform the noisy image into the orthogonal domain by discrete 2D wavelet transforms.
2. Apply hard or soft thresholding to the noisy detail coefficients of the wavelet transform. Choose a shrinkage rule and apply the threshold according to this rule.
3. Perform the inverse discrete wavelet transform (IDWT) to obtain the denoised image. Here, the threshold plays an important role in the denoising process. A small threshold value will retain the noisy coefficients, whereas a large threshold value not.

The work in this chapter investigates denoising of Single Photon Emission Computed Tomography (SPECT) brain images denoising using several methods (see Figure 4.1). We mainly work in the wavelet domain. In Chapter 4 we also describe different methodologies for noise reduction, giving insight into which algorithm should be used to find the most reliable estimate of the original image, given its degraded version. Finally we compare several denoising methods. Each method's assumptions, advantages, and limitations are also discussed in this chapter.

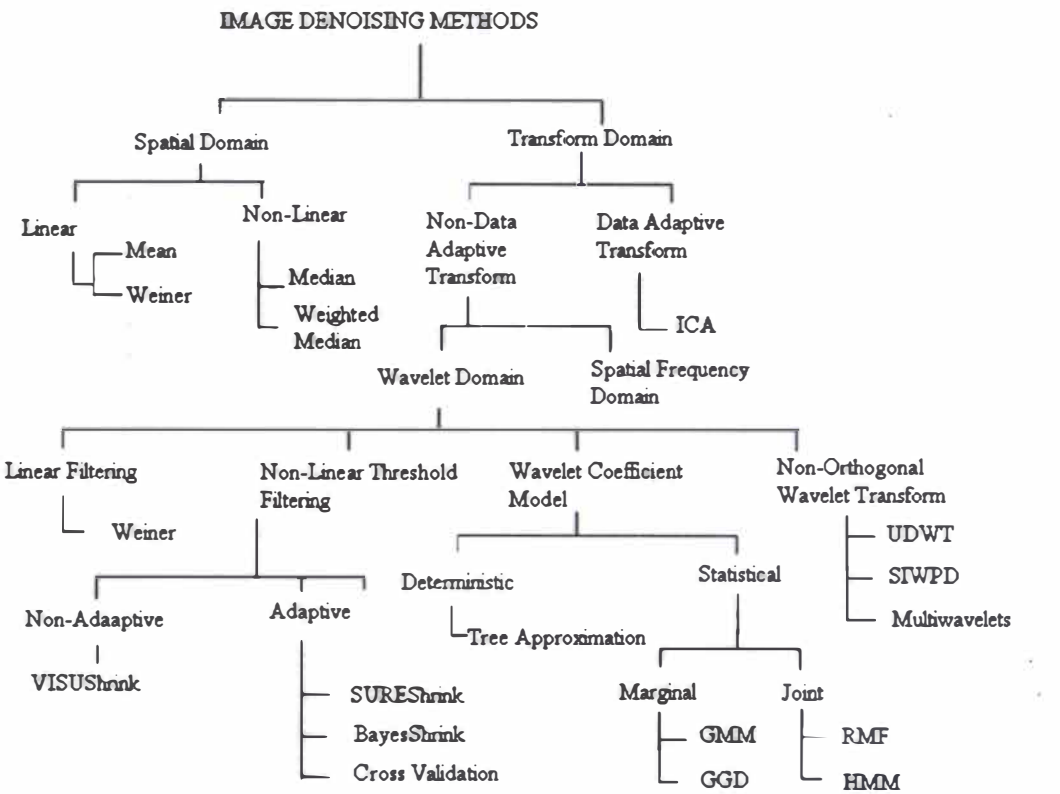


Figure 4.1 Classification of image denoising methods (Motwani, et al., 2004) [κ]

Figure 4.1(Motwani et al., 2004) gives a classification and listing of image denoising methods. In Figure 4.1 RMF refers to Random Markov Field models; HMM to Hidden Markov Models; GMM to the Gaussian Mixture Model; GGD refers to the Generalized Gaussian Distribution. VISU, SURE and BayesShrink are wavelet shrinkage methods (Motwani, et al., 2004).

The SPECT image analysed and its 2D slices are shown in Figure 4.2.). The top left image is coronal with the top (superior) of the head displayed at the top; the left shown on the left (consider the subjectas viewed from behind. The bottom left image is axial with the front (anterior) of the head at the top and the left shown on the left - as if the subject is viewed from above. The top-right image is sagittal with the front (anterior) of the head at the left and the top of the head shown at the top in Figure 4.2 (left) - as if the subject were viewed from the left.

In this Chapter we use the bottom left image with varying transverse slices which are shown in Figure 4.2 (right) as 2D brain images obtained from slice one, which is the bottom of the brain, up to slice 68, which is the top of the brain. Brain images are realigned and normalized into a

[κ] Author has kindly allowed using this figure.

standard brain space so that comparisons can be made of the same area in each brain (Brett et al., 2003). Single photon emission computed tomography (SPECT) has limited spatial resolution and low signal to noise ratio. More details about SPECT analysis of brain image data will be discussed in Chapter 5, in addition to statistical parametric modelling (SPM) and independent component analysis (ICA) (see also Chapter 1).

4.2 Overview Denoising of the Mathematics of Wavelet

Wavelets are becoming an increasingly important tool in image analysis. Spatial frequency domain methods are not optimal for removing noise (Lu et al., 2007; Motwani, et al., 2004). Wavelet methods can better resolve frequencies varying in time. Before applying denoising to the brain image, we provide a review on 2D wavelets theory.

4.2.1 Two-Dimensional Wavelets

The basic theory of wavelets for the one dimension case is given in Chapters 2 and 3. This section describes the construction of two-dimensional wavelets. The extension to more than two dimensions can be accomplished by similar means. We consider analysing two-dimensional signals $f(x,y)$ that are square integrable over the outer product of unidimensional wavelets (Mallt, 1998; Ogden, 1997);

$$\Psi_{j_1,j_2,k_1,k_2}(x,y)=\psi_{j_1,k_1}(x)\psi_{j_2,k_2}(y), \tag{4.1}$$

where denotes $\psi_{j_1,k_1}(x)$ is two dimensional mother wavelet for x with the dilation index j_1 and the translation index k_1.

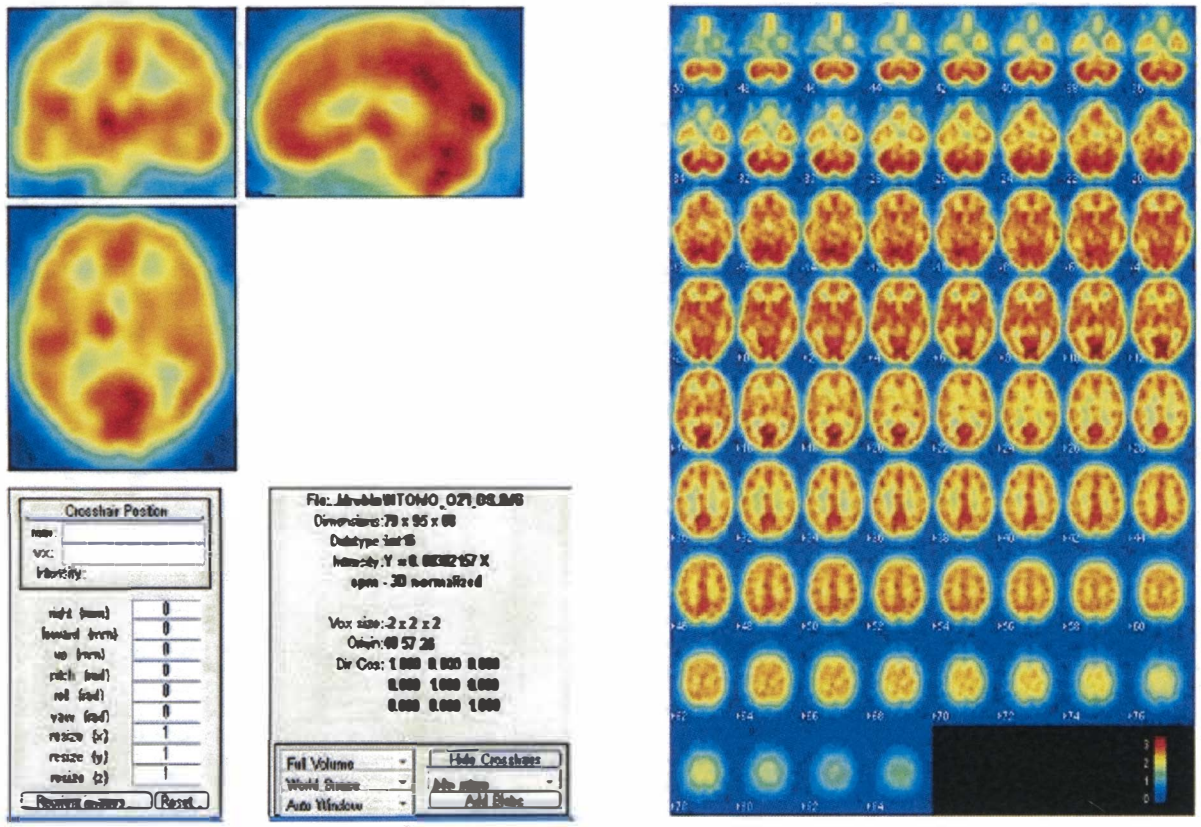


Figure 4.2 The SPECT brain image and the SPECT 2D image for the various slices from the bottom of the brain (the first top slice) to the top of the brain (the last slice)

It is straightforward to show that the Ψ's as defined above are indeed wavelets and that they form an orthonomal basis for $L^2\left(\mathbb{R}^2\right)$ but this approach is too simplistic to retain the optimal features of multiresolution analysis (MRA: see Chapters 2 and 3). Let us begin extending our one-dimensional MRA by defining the two-dimensional space V_0 as the tensor product of two one-dimensional spaces;

$$V_0 = V_0 \otimes V_0 = \overline{\text{span}\left\{f(x)g(y), f, g \in V_0\right\}}. \tag{4.2}$$

Then for any j, this set of a two-dimensional version of the MRA, with the following properties (Mallat, 1998; Ogden, 1997; Motwani, et al., 2004);

1. $\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$
2. $\bigcap_{j\in\mathbb{Z}} V_j = \left\{(0,0)\right\}, \overline{\bigcap_{j\in\mathbb{Z}} V_j} = L^2\left(\mathbb{R}^2\right)$
3. $f \in V_j$ if and only if $f\left(2\bullet, 2\bullet\right) \in V_{j+1}$
4. $f \in V_0$ implies $f\left(\cdot - k_1, \cdot - k_2\right) \in V_0$ for all $k_1, k_2 \in \mathbb{Z}$.

Also, defining $\Phi(x,y) = \phi(x)\phi(y)$ and

$$\Phi_{j,k_1,k_2}(x,y) = 2^j \Phi\left(2^j x - k_1, 2^j y - k_2\right), \tag{4.3}$$

we have the final property,

5. The set $\left\{\Phi_{j,k_1,k_2}, j, k_1, k_2 \in \mathbb{Z}\right\}$ constitutes an orthonormal basis for V_0, given in Equation (4.2).

Notice that this setup differs from the previously mentioned MRA (Chapter 2) by allowing only a single dilation index. To continue this extension into two-dimensions, we now aim to find complimentary spaces that represent the "detail signal" between successive approximations. As in the one-dimensional case, define W_j to be the orthogonal compliment of V_j in V_{j+1}. Breaking down this MRA gives more insight into the nature of the spaces W_j,

$$
\begin{aligned}
V_{j+1} &= V_{j+1} \otimes V_{j+1} \\
&= \left(V_j \oplus W_j\right) \otimes \left(V_j \oplus W_j\right) \\
&= V_j \otimes V_j \oplus \left(\left(V_j \otimes W_j\right) \oplus \left(W_j \otimes V_j\right) \oplus \left(W_j \otimes W_j\right)\right) \\
&= V_j \oplus W_j \ .
\end{aligned}
\tag{4.4}
$$

Thus the detail space W_j is itself made up of three orthogonal subspaces. Bases for these three sub-detail spaces are the corresponding tensor products of their components, so we can define the following two-dimensional wavelets;

$$
\begin{aligned}
\Psi^h(x,y) &= \phi(x)\psi(y); \\
\Psi^v(x,y) &= \psi(x)\phi(y); \\
\Psi^d(x,y) &= \psi(x)\psi(y).
\end{aligned}
\tag{4.5}
$$

Dilating and translating (as in Chapter 2), we have an orthonormal basis for W_j as follows; $\left\{\Psi^m_{j,k_1,k_2}(x,y),\ m=h,v,d;\ k_1,k_2 \in \mathbb{Z}\right\}$ where $\Psi^m_{j,k_1,k_2}(x,y) = 2^j \Psi^m\left(2^j x - k_1, 2^j y - k_2\right)$.

It follows that an orthonormal basis for $L^2\left(\mathbb{R}^2\right)$ $\left\{\Psi^m_{j,k_1,k_2},\ m=h,v,d;\ j,k_1,k_2 \in \mathbb{Z}\right\}$

Mallat (1989) notes that the three sets of wavelets (see Equation (4.5)) correspond to specific spatial orientations. The detail image associated with each of these three orientations will emphasize to edges in the image in the indicated direction. Specifically, the wavelet Ψ^h corresponds to the horizontal direction, the wavelet Ψ^v to the vertical direction, and Ψ^d corresponds to the diagonal direction. The decomposition and reconstruction algorithms for the two-dimensional case are closely related to the corresponding one-dimensional algorithms.

Note that $\phi \in V_0$ and also $\phi \in V_1$ since $V_0 \in V_1$. Since $\left\{\phi_{1,k}, k \in \mathbb{Z}\right\}$ is an orthonormal basis for V_1, there exists a sequence h_k such that the two-scale relationship is $\phi(x) = \sum_{k\in\mathbb{Z}} h_k \phi_{1,k}(x)$ (Ogden, 1997) and that the sequence elements may be written $h_k = \left\langle \phi, \phi_{1,k} \right\rangle$. This sequence h_k is a square-summable sequence and the two-scale relationship $\phi(x) = \sum_{k\in\mathbb{Z}} h_k \phi_{1,k}(x)$, relating functions with differing scaling factors, is also known as the dilation equation. By applying the two-scale relationship $\left(\phi(x) = \sum_{k\in\mathbb{Z}} h_k \phi_{1,k}(x)\right)$ twice and to the two dimensional wavelet in Equation (4.3), we arrive at the two dimensional two-scale relationships (Ogden, 1997):

$$\begin{aligned}\psi_{j,k} &= \sum_{l\in\mathbb{Z}} h_l 2^{j/2}\phi_{1,l}\left(2^j x-k\right)\\ &= \sum_{l\in\mathbb{Z}} h_l 2^{(j+1)/2}\phi\left(2^{j+1}x-2k-l\right)\\ &= \sum_{l\in\mathbb{Z}} h_l\phi_{j+1,l+2k}\left(x\right)\\ &= \sum_{l\in\mathbb{Z}} h_{l-2k}\phi_{j+1,l}\left(x\right).\end{aligned} \tag{4.6}$$

From Equation (4.6) we arrive at the two dimensional two-scale relationship (Ogden, 1997) as follows:

$$\Phi_{j,k_1,k_2}\left(x,y\right)=\sum_{l_1\in\mathbb{Z}}\sum_{l_2\in\mathbb{Z}} h_{l_1-2k_1}h_{l_2-2k_2}\Phi_{j+1,l_1,l_2}\left(x,y\right).$$

An expression relating each scaling function $\phi_{j,k}$ to scaling functions and wavelets at level j-1 can be derived as follows:

$$\phi_{1,k}\left(x\right)=\sum_{l\in\mathbb{Z}} a_{2l-k}\phi_{j-1,l}+b_{2l-k}\psi_{j-1,l} \tag{4.7}$$

where $a_{2k}=\left\langle\phi_{1,0},\phi_{0,k}\right\rangle$ and $a_{2k-1}=\left\langle\phi_{1,1},\phi_{0,k}\right\rangle$, $b_{2k}=\left\langle\phi_{1,0},\psi_{0,k}\right\rangle$ and $b_{2k-1}=\left\langle\phi_{1,1},\psi_{0,k}\right\rangle$.

Applying the result in Equation (4.7) twice gives an analogous result in two dimensions, relating a scaling function, at any level, to scaling functions, and wavelets at a coarser level (Ogden, 1997):

$$\begin{aligned}\Phi_{j,k_1,k_2}\left(x,y\right)=\sum_{l_1\in\mathbb{Z}}\sum_{l_2\in\mathbb{Z}}\Big[&a_{2l_1-k_1}a_{2l_2-k_2}\Phi_{j-1,l_1,l_2}\left(x,y\right)\\ &+a_{2l_1-k_1}a_{2l_2-k_2}\Psi^1_{j-1,l_1,l_2}\left(x,y\right)\\ &+b_{2l_1-k_1}a_{2l_2-k_2}\Psi^2_{j-1,l_1,l_2}\left(x,y\right)\\ &+b_{2l_1-k_1}b_{2l_2-k_2}\Psi^3_{j-1,l_1,l_2}\left(x,y\right)\Big].\end{aligned} \tag{4.8}$$

Naturally, analogous versions of these formulae are applied to give the decomposition and reconstruction algorithms for DWT-MRA coefficients, which can be represented in terms of the high pass and low pass filters (Mallat, 1998; Ogden, 1997) (see Figure 4.3).

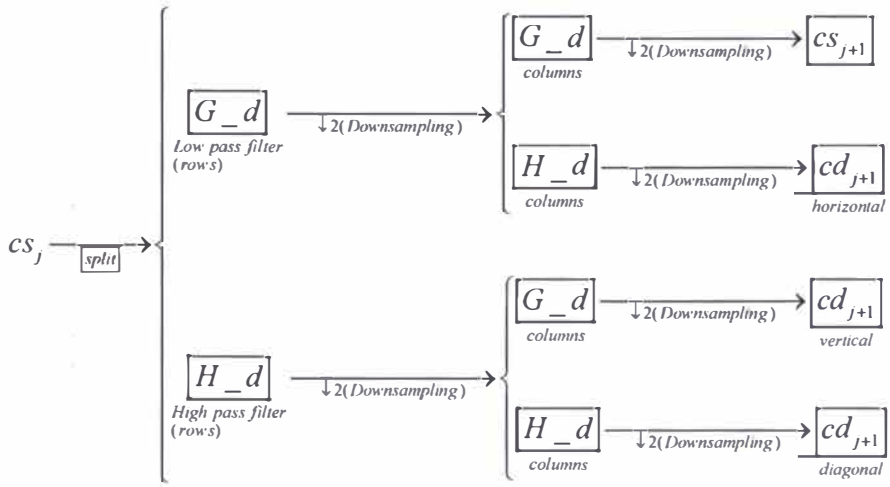


Figure 4.3 Schematic diagram of the two-dimensional wavelet decomposition (G and H denote the low and high pass filters, respectively. $d_j^{(k)}$ denote the details. s_{j-1} is a smoothing of higher level scaling coefficients).

Decomposition is the result of a two-step process, which is represented schematically in Figure 4.3. To begin with, we regard the matrix of scaling function coefficients as a 2D signal, each row being thought of as a separate (1D) signal. The first step consists of applying high filter H for decomposition (H_d) and low filter G for decomposition (G_d) to each row of the matrix, the intermediate results being two matrices with the same number of rows, but half as many columns as the original matrix. Regarding each of these matrices as consisting of columns of (1D) signals, the filters H and G are applied to the columns, giving four final square matrices, each with half as many rows and half as many columns as the original matrix (see the Figure 2.1). The four resultant matrices correspond to the scaling function and the three wavelets. The matrix s_{j-1} is a smoothing of the higher-level scaling coefficients; the matrices $d_{j+1}^h, d_{j+1}^v, d_{j+1}^d$ represent the horizontal, vertical, and diagonal detail components, respectively.

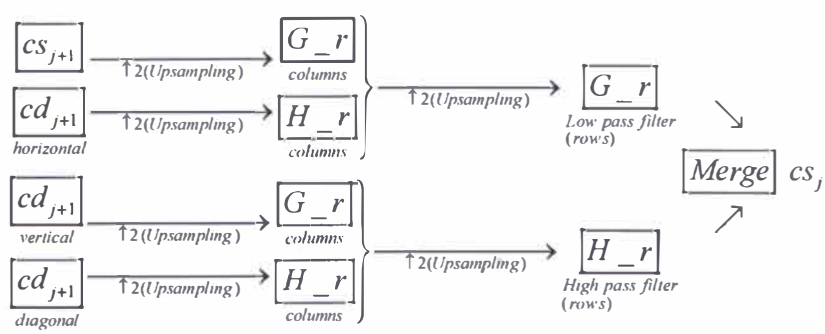


Figure 4.4 Schematic diagram of the two-dimensional inverse DWT reconstruction and (H_r) denotes the high reconstruction filter.

The wavelet decomposition of an image can begin by regarding the original matrix elements as the top-level scaling function coefficients. The matrix of scaling function coefficients, that result from a one-level decomposition, is a smoothing of the original image. The other three matrices represent detail images, each according to its directional orientation. The results of the two-dimensional wavelet decomposition are efficiently stored in the format displayed in Figure 4.4 (see also Figure 2.2), similar to that of Mallat (1989). The first step consists of applying high filter H for reconstruction (H_r) and low filter G for reconstruction (G_r) to each column of the matrix. If the decomposition were to proceed further, the square in the lower left-hand corner of the Figure 4.4 would be replaced by four equally sized squares, the lower left-hand square in turn containing the matrix of scaling function coefficients at level j-1.

4.2.2 Standard Denoising

Whilst noise affects all frequencies, the signal tends to dominate the low-frequency components, so the bulk of the high-frequency components above a specific threshold will be due to noise (Gencay et al., 2002; Ogden 1997).

The DWT can be used as an easy and fast denoising method of a noisy signal. If we take only a limited number of the highest coefficients of the DWT, and perform an inverse transform (IDWT) (with the same wavelet basis) we obtain more or less a denoised signal. There are several ways by which to choose the coefficients that will be kept. Here, only two simple methods were attempted: hard and soft thresholding (Donoho & Johnston, 1994; DeVore & Lucier, 1992). The wavelet denoising procedure used consists of the following steps (Gencay et al., 2002)

1. Compute a partial DWT (PDWT) using Equations (2.4) and (2.5) with level J_p: It is mathematically formulated as follows;

$$\begin{pmatrix} v_{11} & v_{12} & \cdots & v_{1n} \\ v_{21} & v_{21} & \cdots & v_{2n} \\ & & \cdots & \\ v_{n1} & v_{n1} & \cdots & v_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} \Rightarrow \sqrt{n} \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

where $(\eta_1 \ \eta_2 \cdots \eta_n)^T$ are the decomposed noisy signals.

2. Select a threshold η. The threshold may be a scalar or vector and may be level dependent or independent. The standard deviation of the noise, if known, may be estimated using Equation (4.9). The most commonly used estimator of σ_ε which is the maximum absolute deviation (MAD) standard deviation which uses only the first scale wavelet coefficients W_1, defined as

$$\hat{\sigma}_{MAD} = \frac{\text{median}\left(|W_{1,0}|, |W_{1,1}|, \ldots, |W_{1,N/2-1}|\right)}{0.675}. \qquad (4.9)$$

The denominator in Equation (4.9) is needed so as to rescale the numerator of Equation (4.9) so that $\hat{\sigma}_{MAD}$ is tuned to estimate the standard deviation for Gaussian white noise. The median is used in order to produce a robust estimator of the noise variance, isolating it from the signal variance (Johnstone & Silverman, 1997; Gencay et al., 2002; Ogden, 1997).

3. Put the wavelet coefficients $W_{j,t}$, $j = 1, \ldots, J_p$ and $t = 0, \ldots, N/2^j - 1$, through a thresholding rule using a threshold η to produce a collection of thresholded wavelet coefficients $\widetilde{W}_p$. Schematically we have,

$$\begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \Rightarrow \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix}$$

where $(\zeta_1\ \zeta_2 \cdots \zeta_n)^T$ is the result of filtering out the noise from the decomposed noisy signals, which is obtained by

$$\zeta_i = \begin{cases} \eta_i & \text{if } |\eta_i| \geq \lambda_{n,\delta} \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } i = 1, \cdots, n.$$

4. Produce an estimate of X by performing the inverse partial DWT $\widetilde{X} = W_p^T \widetilde{\mathcal{W}}_p$.

Figure 4.5 shows all the steps schematically. Wavelet denoising reduces high frequency noise and a wavelet transform is computed from the original image. Then the noise level at each wavelet scale is estimated separately. This defines a threshold for zeroing the wavelet coefficients. Other wavelet coefficients are shrunk according to local variance estimation (Donoho, 1995). After applying the inverse wavelet transform (IDWT), the image is renormalized (Figure 4.5).

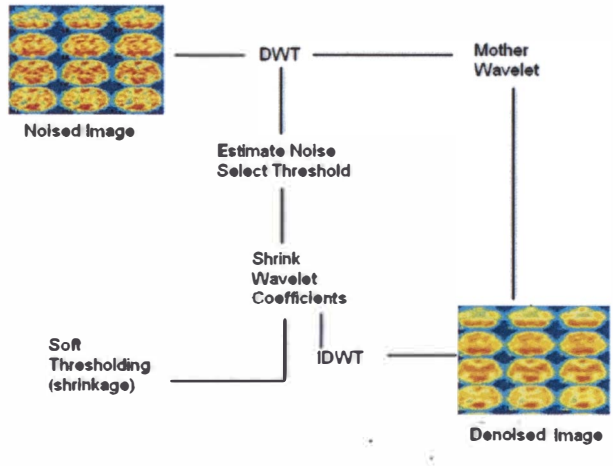


Figure 4.5 Flow diagram illustrating wavelet denoising

The 2D DWT described above is used in many applications, such as, compression, denoising and watermarking applications (Mahmood & Selin, 2006). The 2D DWT is built with separable orthogonal mother wavelets, which possess acertain regularity. At every iteration of the DWT, the lines of the input image are first low-pass filtered (high frequency) with a filter associated with the impulse response G and a high-pass filter (low frequency) denoted by H. Next the lines of the two images (as obtained at the output of the two filters) are decimated by a factor of 2. Subsequently, the columns of the two images are low pass filtered with G and high-pass filtered with H. The columns of the resultant four images are additionally decimated (again by a factor of 2). Four new sub-images are thereby generated. The first, obtained after two low-pass filtering

procedures, is called the approximation (or smoothing) sub-image (or LL image). The other three are designated as the detail sub-images: LH (horizontal detail), HL (vertical detail) and HH (diagonal detail) (see Figure 4.6). The LL image then represents the input for the next iteration. In what follows, the coefficients of the DWT will be denoted by D_m^k, where m represents the iteration index (the resolution level) and k=1, for the HH image, k=2, for the HL image, k=3, for the LH image and k=4, for the LL image. These coefficients are calculated using the following relationship:

$$D_m^k(p,q) = \left\langle x(\tau_1,\tau_2),\ \psi_{m,p,q}^k(\tau_1,\tau_2) \right\rangle \tag{4.10}$$

where the wavelets can be factorised:

$$\psi_{m,p,q}^k(\tau_1,\tau_2) = \alpha_{m,p,q}^k(\tau_1)\ \beta_{m,p,q}^k(\tau_2) \tag{4.11}$$

and the two factors can be computed using the scale function $\phi(\tau)$ and the mother wavelet $\psi(\tau)$ with the aid of the following relation:

$$\alpha_{m,p,q}^k(\tau) = \begin{cases} \phi_{m,p}(\tau), k = 1,4 \\ \psi_{m,p}(\tau), k = 2,3 \end{cases} \tag{4.12}$$

$$\beta_{m,p,q}^k(\tau) = \begin{cases} \phi_{m,p}(\tau), k = 2,4 \\ \psi_{m,p}(\tau), k = 1,3 \end{cases} \tag{4.13}$$

where

$$\phi_{m,p}(\tau) = 2^{-m/2}\phi\left(2^{-m}\tau - p\right) \tag{4.14}$$

$$\psi_{m,p}(\tau) = 2^{-m/2}\psi\left(2^{-m}\tau - p\right). \tag{4.15}$$

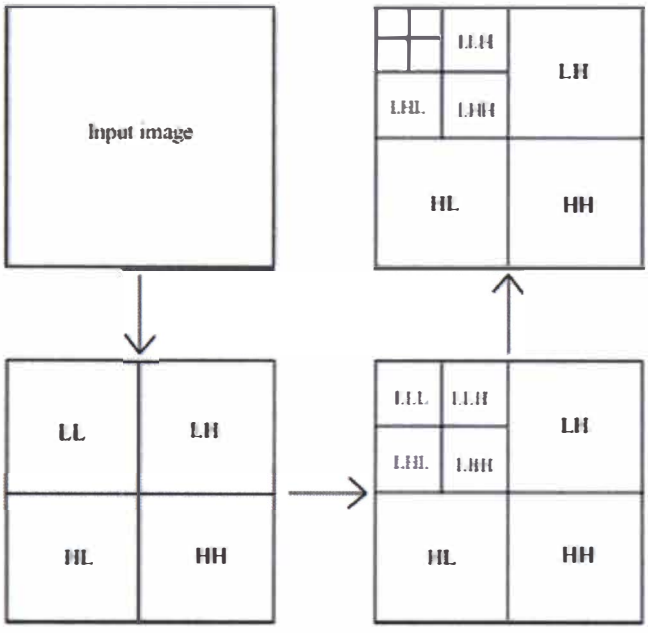


Figure 4.6 2D-DWT with 3 level decomposition.

4.2.3 Two-dimensional DWT of Images

The discrete wavelet transform (DWT) of image signals produces a non-redundant image representation, which provides superior spectral localization of image formation, compared with other multi scale representations (Mohideen et al., 2008). The DWT can be interpreted as signal decomposition into a set of independent, spatially oriented frequency channels. The signal is passed through two complementary filters and emerges as two signals, approximation and details. This is called decomposition. The components can be assembled back into the original signal without loss of information. This process is called reconstruction. The mathematical manipulation involves the discrete wavelet transform (DWT) and the inverse discrete wavelet transform (IDWT). An image can be decomposed into a sequence of different spatial resolution images using DWT. In the case of a 2D image, an N level decomposition can be performed resulting in $3N+1$ different frequency bands namely: LL, LH, HL and HH as shown in Figure 4.6.

The next level of wavelet transform is applied to the low frequency sub-band image LL only. The Gaussian noise will be nearly averaged out in low frequency wavelet coefficients. Therefore, only the wavelet coefficients in the high frequency levels need to be thresholded. Using Equation (4.2), we perform a 2D-DWT, which is shown in Figure 4.7.

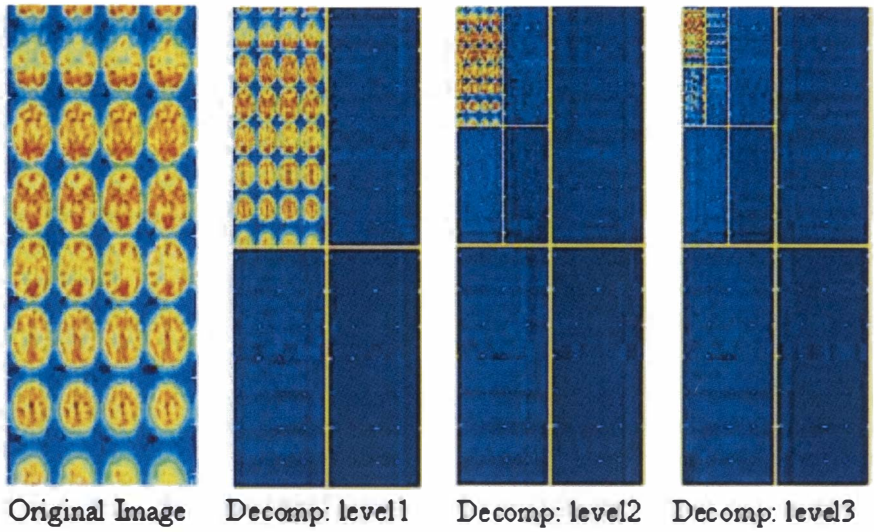


Figure 4.7 2D-DWT with 3 levels of decomposition for Patient 28 (P28).

In this study we apply non-linear denoising through thresholding given its ability to adapt to rapidly changing nonstationary features in an observed process (Donoho & Johnson, 1998). The simplest method of non-linear wavelet denoising is via thresholding denoising (see the theory of thresholding in Section 3.2.1). Figures 4.8- 4.10 are given below which inform on both the horizontal, diagonal and vertical details in addition to the level j (j= 1, 2, 3) used. Figures 4.8-4.10 are arranged so as to emphasise the tree structure of the 2D-DWT decomposition.

Our procedure is as follows: [1] the image is transformed into the orthogonal domain by the DWT; [2] the detail wavelet coefficients are modified according to the shrinkage algorithm used; [3] finally the inverse wavelet is used to reconstruct the denoised image. Matlab Wavelet Toolbox (Misiti et al., 1997) is utilized to obtain the DWT of the inputted image. In each sub-band, individual pixels of the image are shrunk based on the threshold selection. Shrinking pixels creates the resultant denoised DWT. The inverse DWT (IDWT) is the resultant denoised image.

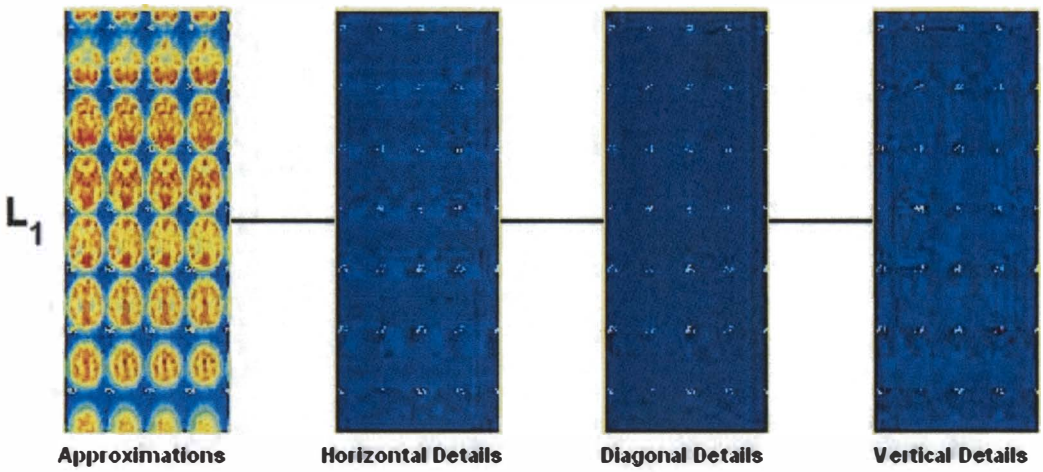


Figure 4.8 Decomposition level 1 for P28.

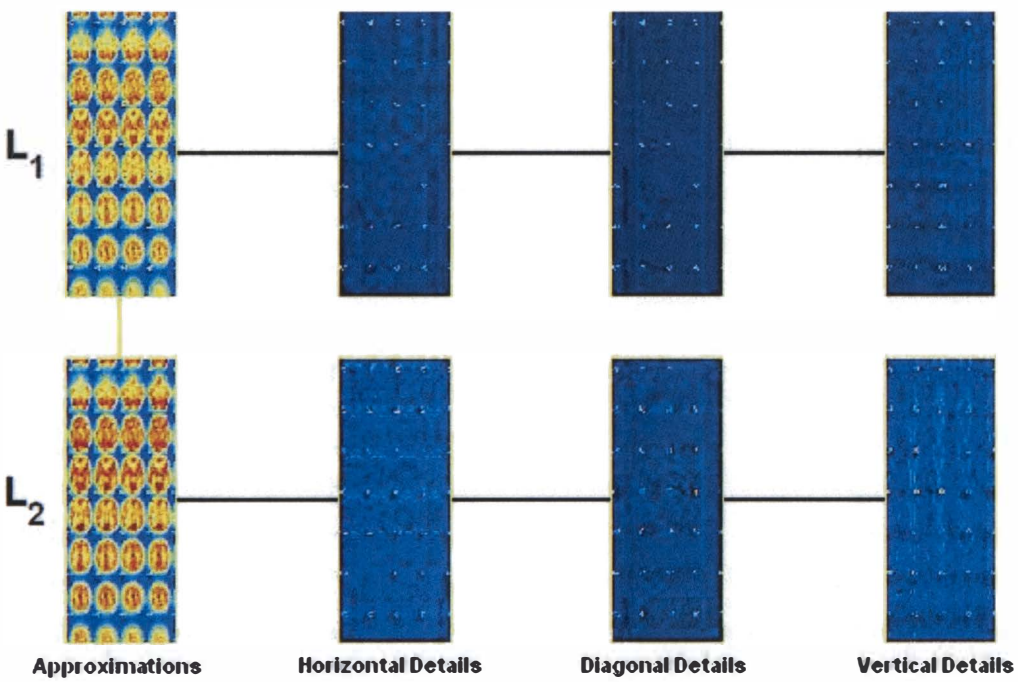


Figure 4.9 Decomposition up to level 2 (L_1, L_2) for P28.

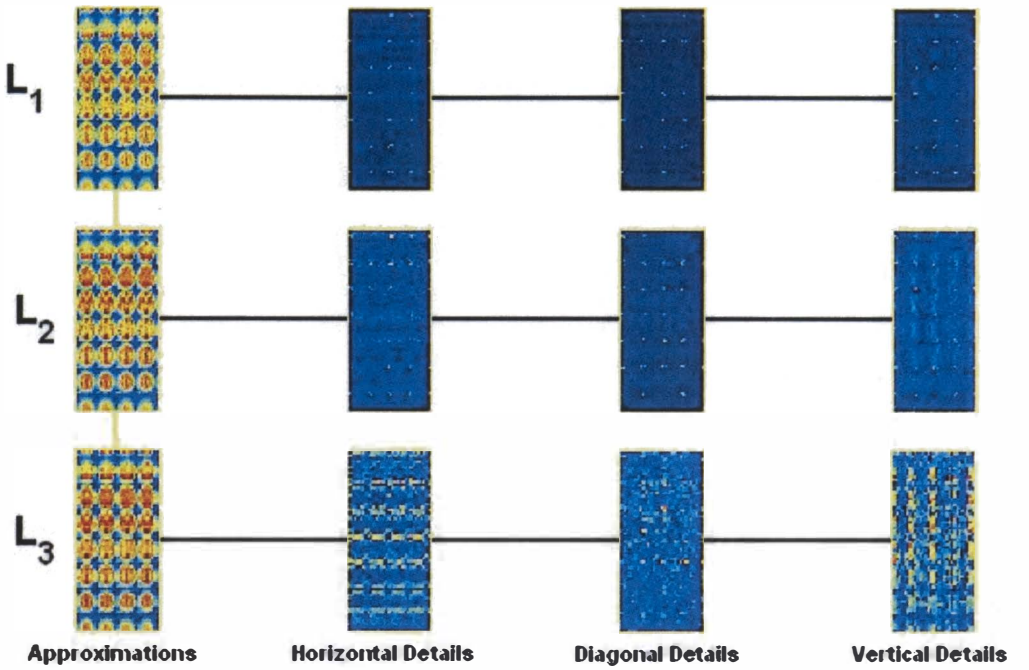


Figure 4.10 Decomposition up to level 3 (L_1, L_2, L_3) for P28.

We examine four different wavelet types namely the Haar, Daub 2, Daub 4 and LaDaub 8 (see Table 4.2).

Choosing a very large threshold will make it difficult for a coefficient to be judged significant and so be included in the reconstruction, resulting in an oversmoothed estimate. On the other hand, a very small threshold value will allow many coefficients to be included in the reconstruction resulting in an undersmoothed estimate. The key parameter in all thresholding rules is η (Gencay et al., 2002). Optimal thresholding occurs when the thresholding parameter is set to the noise level $\eta = \sigma_\varepsilon$. Thus, the coefficients $\boldsymbol{o}_t = \mathcal{W}y = \mathcal{W}x + \mathcal{W}\varepsilon = s_t + \varepsilon'$, where $t = 1, \ldots, N$, and $\mathcal{W}$ is an $N \times N$ orthonormal matrix, $y_t = x_t + \varepsilon$, $t = 1, \ldots, N$, x_t is a time varying signal on equally spaced points with $\varepsilon \sim N(0, \sigma_\varepsilon^2)$, to be included in the regression model, satisfy the hard thresholding rule, with λ being set to the known noise level that is $\delta_{\sigma_\varepsilon}^{H}(\boldsymbol{o}_t)$. Setting $\eta < \sigma_\varepsilon$ will allow unwanted noise to enter the estimate, while setting $\eta > \sigma_\varepsilon$, will destroy information that really belongs to the underlying signal. Setting $\eta = \sigma_\varepsilon$ is only possible if we know σ_ε (Donoho & Johnstone, 1994). The ideal risk of the diagonal projection (dp) associated this oracle knowledge is given by

$$R(\mathrm{dp}, o) = \sum_{i=1}^{N} \min\left(|s_i|, \sigma_\varepsilon\right), \tag{4.16}$$

Actually this risk cannot be attained in practice by any estimator (Donoho & Johnstone, 1994, 1995a, 1995b; Ogden, 1997; Gencay et al., 2002; Nason, 2008). We propose using a thresholding rule that attempts to come close to the risk R(dp, *o*).

In this chapter we use *Sureshrink* (Donoho & Johnstone, 1995b) to denoise the image (see Section 4.3).

4.3 Results and Discussion

In this study we compare four different denoising methods Tree-Adapted wavelet shrinkage (TAWS), *SureShrink*, Median Adaptive and Median (see Table 4.1)) via three error measures (Sup, RMS and SNR) (Table 4.1). Secondly we design two different wavelet transform schemes the DWT and MODWT (see Chapter 2) with four different wavelet bases (Haar, Daub 2, Daub 4, Ladaub 8) (Table 4.2) and test them at various window sizes for denoising performance (Table 4.3). The aim is to study the suitability of different wavelet bases and also of different window sizes for denoising brain image data discussed in detail in Chapter 5. Thirdly in Section 4.3.3 we compare wavelet denoising versus independent component analysis (ICA) denoising (using P28 brain image data).

4.3.1 Comparison of the four different denoising methods

We examined four different denoising methods by comparing error estimates obtained via FAWAVE (version 2.0, Walker, 2008). The first method is the Tree- Adapted Wavelet Shrinkage (TAWS) technique implemented by the TAWS algorithmas described in Walker (1999; 2000; 2002; 2003) and in Walker & Chen (2000). In the TAWS procedure, shrinkage, as well as averaging, is implemented which reduces the computation time required. The second routine is the famous *SureShrink* method (Donoho & Johnson, 1995b, 1994; Gencay, et al.,2002; Ogden, 1997). *SureShrink* uses a hybrid of the universal threshold (Donoho & Johnson, 1994, 1995b) and the Stein's Unbiased Risk Estimator (SURE) threshold (Donoho & Johnson, 1994, 1995b; Gencay et al., 2002). We now describe our remaining non-linear methods (Median and Adapted median, see Table 4.1).

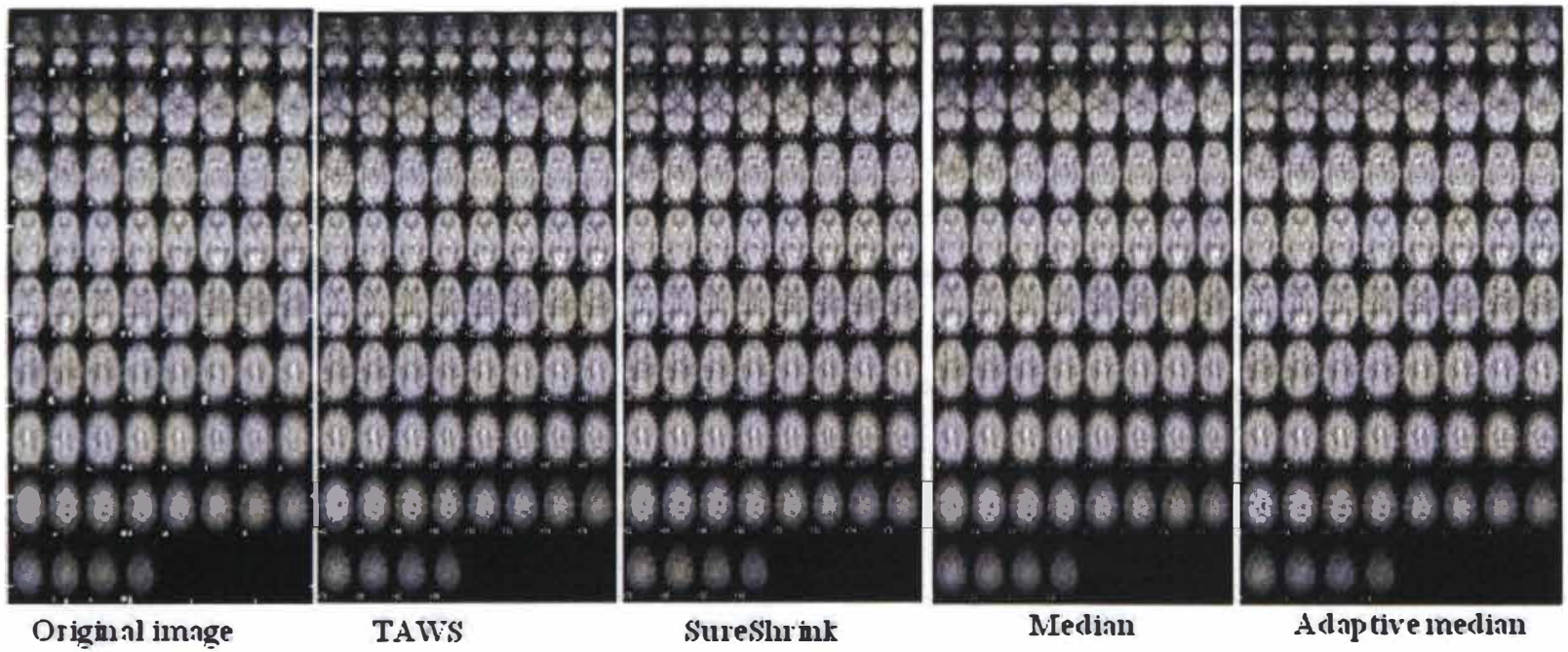


Figure 4.11 Denoised images using several denoising methods for P28.

With non-linear filters in spatial filtering, noise is removed without any attempt to explicitly identify it. Spatial filters employ a low pass filtering on groups of pixels with the assumption that noise occupies the higher region of the spectrum. In recent years, a variety of non-linear median type filters have been developed to overcome the blurring of images (Ben Hamza et al., 1999; Hardie & Barner, 1994; Yang et al., 1995). The third method we use is non-linear filtering which is a 3 by 3 median denoising (Fitch et al., 1984). This is used to denoise images that are corrupted by impulse noise. The fourth method denoises an image by a 3 by 3 adaptive median filter (Wang & R.A, 1995; Xing & Wang, 2001). In this procedure the pixel value centred on a 3 by 3 region is only replaced by the median, if its values differ from the median by more than $\pm 10\%$. It usually results in a more sharply resolved denoising than does median denoising. Denoised images (P28) via the four methods are given in Figure 4.11. It is difficult to distinguish between the original image and the four denoised images. We then use three ways of measuring the amount of error between the original image and the resultant noisy image. All of these measures aim to provide quantitative evidence for the effectiveness of the given noise removal method (as given in Table 4.1) (Sup, RMS and SNR).

The first error measure is the supreme error (*Sup*), which is taken to be the largest magnitude error between the original image and denoised image values at all pixels (Walker, 1999). The second measure is the root mean square error (*RMS*), which is the classical error between the original image and the denoised image (Walker, 2000). A third measure, commonly used in image processing, is the signal-to-noise ratio (*SNR*) (Walker, 1999, 2000). The results of applying these three measures to P28 are summarized in Table 4.1. When we compare the *Sup* error and *RMS*, across denoising methods, *SureShrink* gives the smallest error. *SureShrink* appears to be superior among the four denoising methods. For the *SureShrink* denoising method the RMS error is 0.246, which is less than the RMS error of 0.565 for TAWS denoising. Both the Median and the Adaptive Median denoising methods are not much different from each other across all error estimate types. Unlike other measures, an increase in SNR represents a decrease in error. The *SureShrink* denoising method appears to be superior to the other three denoising methods according to SNR.

The measures of error discussed above have all been used for many years. Their deficiencies in relation to accurate quantification of the perceptions of our visual systems are well known. It is generally recognized that they have remained in use, despite their deficiencies, mainly because they fit well into the type of mathematics used in image processing, making theoretical predictions concerning their values relatively easy to obtain (Walker, 1999, 2000).

Table 4.1 Error measurements for the SPECT 2D images.

Denoising method	*Sup*	*RMS*	*SNR*
TAWS	5	0.565	48.0941
SureShrink	1	0.246	55.297
Median	254	16.284	18.915
Adaptive Median	254	16.169	18.976

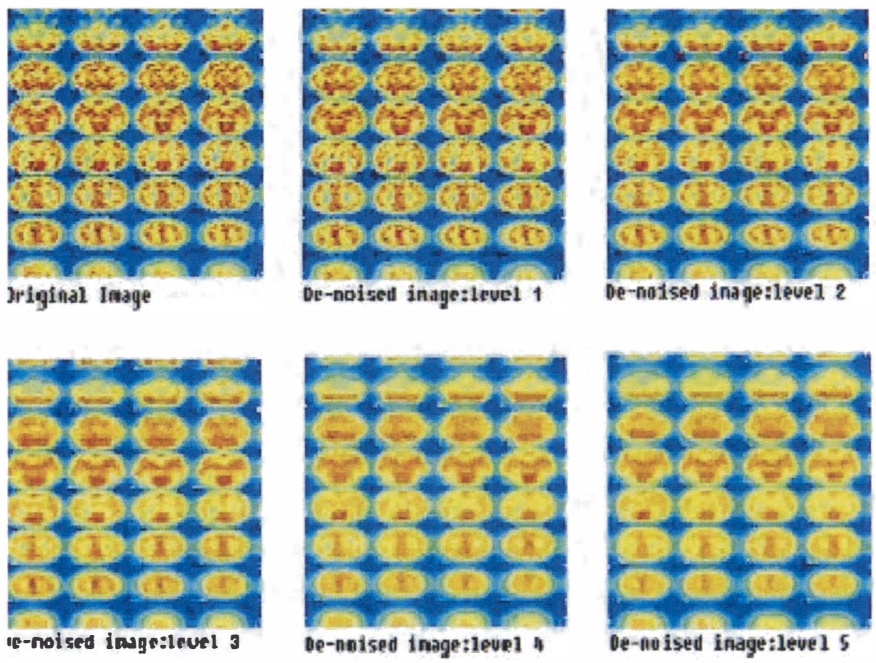


Figure 4.12 Original brain image and denoised images up to 5 levels for P28 using the soft threshold (via the Haar wavelet and DWT).

4.3.2 Wavelet domain for image denoising

Now we consider only the wavelet domain for image denoising. By choosing a threshold that is a sufficiently large multiple of the standard deviation of the random noise, it is possible to remove most of the noise by thresholding the wavelet transform values. We examine how this procedure performs on the P28's brain image using the soft thresholding technique. The soft thresholding method will be compared with hard thresholding.

Our first example of image denoising is examined by using the Haar wavelet on the noisy version of P28's brain image. The result is shown in Figure 4.12. The original image is transformed to various window sizes using the DWT and MODWT.

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components as illustrated by

Figure 4.12. In Figure 4.12, the levels 3, 4, and 5 seem to be oversmoothed and the soft threshold seems to be too aggressive. Nevertheless, the histogram of the details is quite good for level 3 since it is close to a Gaussian distribution (see Figure 4.13).

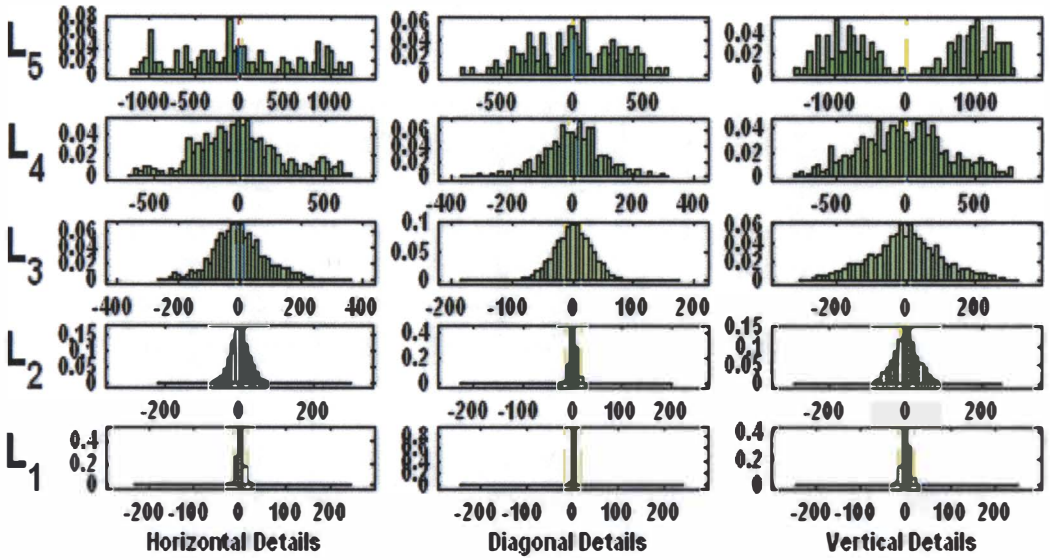


Figure 4.13 Histogram of 5 level details for P28 via the DWT.

Table 4.2 Wavelet soft threshold values used to generate the denoised image of P28 using four different wavelets.

	Wavelet types			
Level	Haar	Daub 2	Daub 4	LaDaub 8
Level 1	16.9	10.3	8.345	7.495
Level 2	15.73	9.607	7.804	7.043
Level 3	14.47	8.867	7.256	6.611
Level 4	13.09	8.086	6.702	6.226
Level 5	11.55	7.274	6.185	5.9

Table 4.3 The standard deviation of the residuals of the denoised image for Patient 28 (P28).

Wavelet transform	Wavelet	Window size					
		256x256	256x512	256x1024	512x256	512x512	512x1024
Level 3							
DWT	Haar	8.2	8.2	5.4	8.2	8.2	5.4
	Daub2	5.1	5.0	3.1	4.9	4.8	3.0
	Daub4	4.2	4.0	2.5	4.1	4.0	2.6
	LaDaub8	3.8	3.7	2.4	3.6	3.5	2.3
MODWT	Haar	14.7	14.9	9.8	14.5	14.7	9.9
	Daub2	11.4	11.4	7.9	11.4	11.4	7.8
	Daub4	10.7	10.7	7.9	10.8	10.8	7.8
	LaDaub8	10.3	10.3	7.9	10.4	10.4	7.8
Level 5							
DWT	Haar	8.3	8.2	5.4	8.4	8.3	5.5
	Daub2	5.2	5.0	3.2	5.0	4.8	3.1
	Daub4	4.2	4.1	2.5	4.2	4.1	2.5
	LaDaub8	3.9	3.8	2.4	3.6	3.6	2.3
MODWT	Haar	17.2	13.4	11.4	16.9	17.1	11.2
	Daub2	12.7	12.6	8.6	12.6	12.4	8.3
	Daub4	11.7	11.6	8.5	11.7	11.6	8.3
	LaDaub8	11.1	11.1	8.4	11.1	11.0	8.3

Given the Gaussianity of level 3 details (Figure 4.13) we now deal only with level 3 wavelets. Next we examine to denoised images using first the DWT and then MODWT using the same choice as before of wavelet bases. We use threshold values for the soft threshold, as given, in Table 4.2.

When we compare Figure 4.14 with Figure 4.15, whilst it is difficult to distinguish in the printed versions between the original image and the four denoised images, Figure 4.15 is more clear with the MODWT denoised images appear to be slightly superior. The residuals between the original image and denoised image confirm this subjective statement (see the standard deviation of the residuals given in Table 4.3).

Note that if the filter cleans up the given image effectively, the residual image is essentially just noise. Therefore small standard deviation (SD) gives a better denoised image (Table 4.3). Comparison with the performance of the wavelet transform was made. Results show that the MODWT produce better results, although the standard deviation of the residuals is often higher than the DWT error.

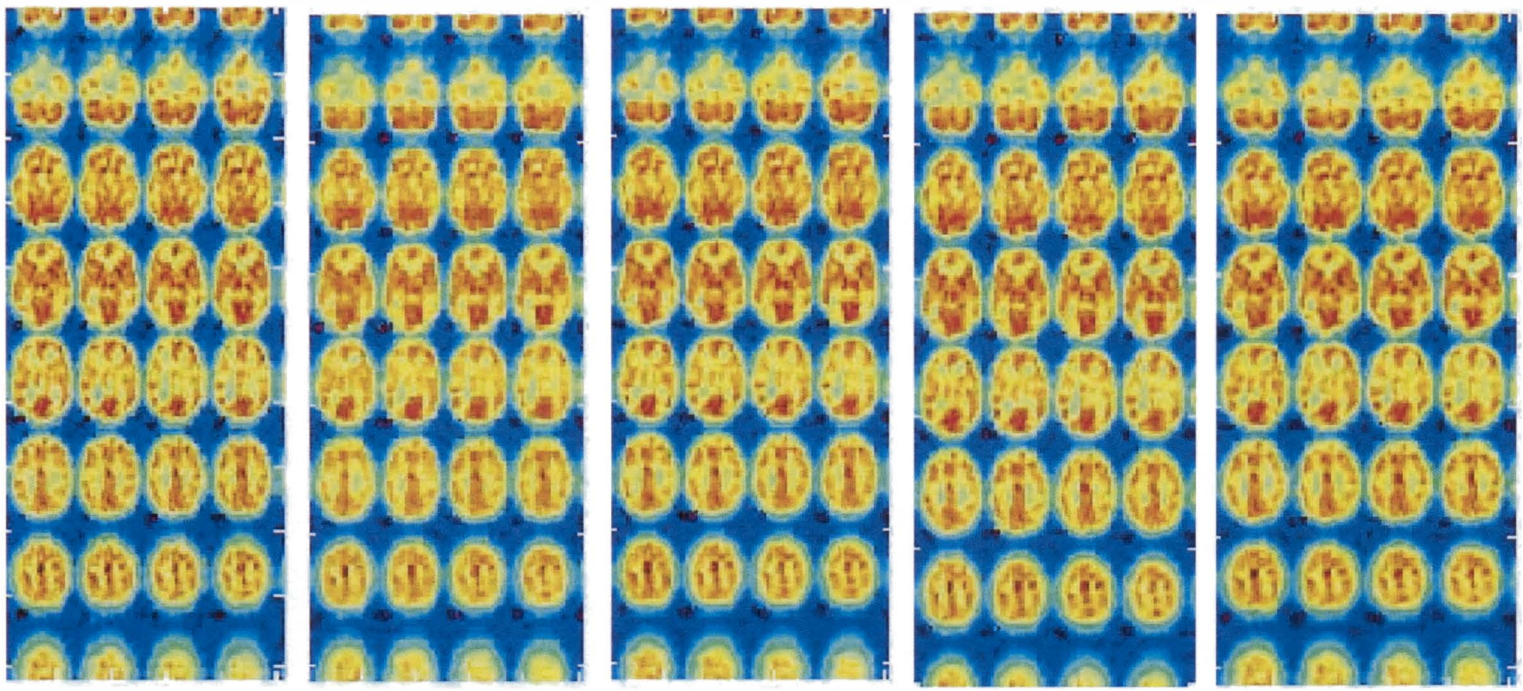


Figure 4.14 Denoised images using the DWT with varying wavelet types (P28).

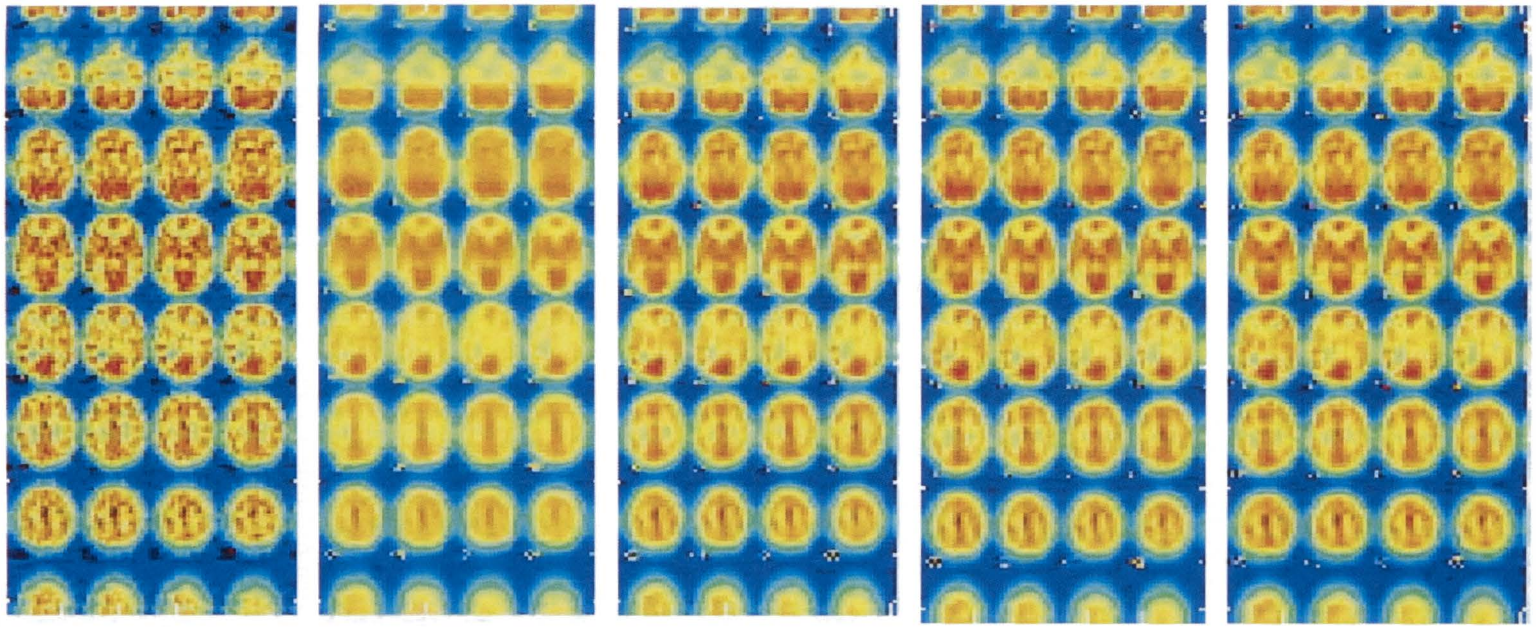


Figure 4.15 Denoised images using the MODWT with varying wavelet types (P28).

When comparing the four different wavelet bases (see Table 4.3), LaDaub8 gives a superior denoised image. We can conclude from our experiment that choice of the wavelet basis is crucial. The experiments were done using a window size of 256x256, 256x512, 256x1024, 512x256, 512x512, and 512x10204 (Table 4.3). The extended images 256x1024 and 512x10204 are good choices. It is noteworthy that the small SD values of residuals do not always correspond to good visual quality.

4.3.3 Comparison between Wavelet and ICA denoising

In Section 4.3.1 we compared the different denoising methods on the basis of three well-defined criteria: the supreme error (*Sup*), the root mean square error (*RMS*), and the signal to noise ratio (*SNR*). Every criterion measures a different aspect of the denoising method. It is easy to show that only one criterion is not sufficient to judge the image, and so one expects a good solution to have a high performance under the three criteria.

In this section we compare two source signal extraction algorithms, namely the Wavelet Denoising (WD) (by soft thresholding) method (using the Haar wavelet) and Independent Component Analysis (ICA) on a patient 28's brain image. Patient 28-2D data denotes the30th slice generated in 3D (79x95x68) among 68 slices of the patient's brain image (P28).

A common definition of SNR is the ratio of mean to standard deviaation of a signal in statistics (Säckinger, 2005; Schroeder, 2000)

$$SNR = \frac{\mu}{\sigma},$$

where μ is the signal mean and σ is the standard deviation of the noise. The higher the ratio, the less obtrusive the background noise is. Recall that the discrete wavelet transform (DWT) is a time-scale representation technique of a signal with a mother wavelet function. Wavelet transforms can thus be used to reduce the noise in a signal by a method called wavelet shrinkage proposed by Donoho (1995) that is mentioned in Chapter 3. DWT localizes information of the deterministic signal into a limited number of the wavelet coefficients (see Equations (2.1)-(2.5) and Figures (2.5)- (2.9) in Chapter 2).

The use of ICA denoising is motivated by the fact that the linear model consists of a set of independent signals additively combined (see Chapters 1 and 5). Since ICA is capable of identifying the components of a mixture, the demixing becomes in effect a denoising procedure (Jung et al., 2000; Makeig et al., 1999; Ylipaavalniemi et al., 2006). To perform ICA, two conditions need to be fulfilled; firstly the components need to be statistically independent and secondly the mixing is required to be linear.

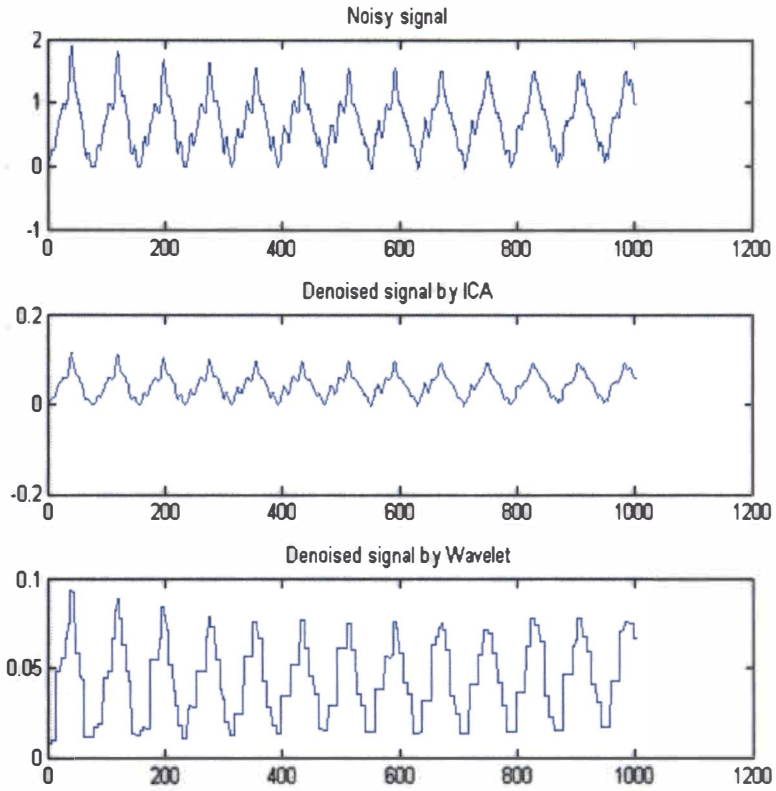


Figure 4.16 Top- Noisy signal for P28-30th slice, the middle is the denoised signal using ICA and the bottom is the denoised signal using Haar wavelet thresholds.

FastICA (Hyvärinen, 2001a; Hyvärinen et al., 2001b) does not differentiate between signal and noise, but basically separates out the components (Bobin et al., 2007; Hyvärinen et al., 2001b; Üzümcü et al., 2003), in contrast to the wavelet denoising (WD) technique. The original image becomes the denoised version of the signal via the use of the Haar wavelet. An application of the FastICA algorithm yields the estimated signal waveform as shown in Figure 4.16, which displays only parts of the signal, given the excessive length of the whole series. We use two different performance measures, namely the signal-to-noise (SNR) ratio and the correlation coefficient, to quantify ICA's capability of recovering the signal. The SNR values are computed both prior to and post ICA denoising (see Table 4.4 for patient 28 (P28) results). The correlation values provide an estimate of the degree of similarity between the original signal and its corrupted variant. The correlation coefficient thus produces a recovery performance measure for the denoising process. Values of SNR and the correlation coefficients for P28 relating the noisy and denoised signals are given in Table 4.4.

Table 4.4 SNR and Correlation coefficients between the noisy and denoised signal: P28.

	SNR	Correlation coefficient
Wavelet Denoising	1.4319	0.851
ICA denoising	1.4869	1.000

SNR is widely used in the signal processing. Commonly mean squared error (MSE) or peak signal to noise ratio (PSNR) are used to compare the performance of denoising method to other because it is simple and easy to compute. Actually the values SNR is smaller than PSNR due to PSNR based on the root of mean square error of the reconstructed image.
In Table 4.4, the values of SNR between WD and ICA are not much different. The actual value of SNR is not meaningful, but the comparison between the two values of different reconstructed images gives a measure of quality. From Table 4.4, we conclude that the performance of ICA remains near optimal while WD degrades substantially. This indicates that ICA is a robust denoising technique, in that its performance is not impacted by the severity of the mixing conditions. Note that the ICA correlation coefficient has a value of 1.0, which corresponds to near perfect waveform matching.. On the other hand, wavelet denoising has a correction coefficient of 0.8581 that shows a poorer performance compared to ICA. Therefore we see that, in the presence of non-Gaussian noise contamination, as in the case of brain image slice signals, WD fails to retrieve the original source signal. In contrast, ICA denoising has a robust and near optimal performance autonomous of the severity of mixing and of the statistical distribution assumed for the mixing components. Additionally ICA does not require outliers to be eliminated (Emir et al., 2003; Matthews et al., 2008; Vigneron et al., 2003).

An application ICA to brain image denoising is described in detail in Chapter 5.

4.4 Conclusions

According to the first part of our study comparing four denoising methods, the *SureShrink* denoising method appears to be slightly superior. We have designed two different wavelet transform schemes, namely the DWT and MODWT with four different wavelet bases and tested their performance for image denoising. The aim of this design was to study the suitability of different wavelet bases and also of different window sizes. In the third part of this chapter we compared wavelet denoising versus ICA. We concluded our discussion of wavelet based denoising of images by using several different wavelet bases for both the DWT and MODWT. The LaDaub8 base performs well in image denoising. Depending on the type of images and the scale of the objects in them one may, however, prefer different wavelet transforms.

In this work, denoising of images was restricted to 2D. However the analysed P28 SPECT brain image is actually 3D data. Denoising is generally performed in 3D space to take advantage of better separation of noise and signal in higher dimensions and the availability of volumetric features available in true 3D data sets.

We conclude that wavelet denoising (WD) shows inferior performance to ICA and WD fails to recover the original source signal since the noise model was not satisfied.
In Chapter 5, independent component analysis (ICA) is employed to remove noise from the 3D SPECT brain image data (this is sourced from the data reported by Turner et al., 2003). ICA methods use a transform, which is estimated from the data itself, in contrast to wavelet based denoising techniques; and as such it is anticipated that ICA may be superior to wavelets, at least in some applications (see Chapter 5).

Chapter 5

5 Independent Component Analysis (ICA) and Statistical Parametric Mapping (SPM): modelling brain function and personality

5.1 Introduction

Independent component analysis (ICA) has been applied to functional magnetic resonance imaging (fMRI) data as an exploratory data analytic technique which aims to discover independently distributed spatial patterns that illustrate underlying source processes in the data (Friston, 1998; McKeown et al., 1998b; McKeown & Sejnowski, 1998c; Van de Ven et al., 2004; Ylipaavalniemi et al., 2006).

The basic goal of ICA (Hyvärinen, 2001a; Lewicki & Olshausen, 1998a; Lewicki & Sejnowski, 1998b; Roberts & Everson, 2001; Stone, 2004) is to find a solution to the Blind Source Separation (BSS) problem by expressing a set of random variables (observations) as linear combinations of statistically independent latent component variables (so-called source signals) (Cardoso, 1991; Lewicki & Olshausen, 1998a; Moulines et al., 1997). By formulation of the ICA model, the data variables are assumed to be linear mixtures of various unknown latent variables, and the mixing system is also unknown. As such ICA is a statistical and computational technique for revealing hidden (latent) factors that underlie sets of random variables, measurements, or multivariate signals (Hyvärinen et al., 2001b). These latent variables are assumed to be nongaussian and mutually independent.

ICA is a much more powerful technique than conventional multivariate methods such as principal components analysis (PCA) or factor analysis (FA), and is capable of finding the underlying factors or sources when these conventional techniques completely fail. ICA is also very successful in finding artefacts (Jung et al., 2001; Vig'ario et al., 2000) but to date less successful in finding components related to brain activity (Hyvärinen et al., 2009). The data analysed by ICA can originate from many different kinds of applications, such as psychometrics, image methods, document databases and econometric studies. Non-invasive methods developed recently for investigating brain function have permitted functional brain function, rather than structural information, to be modelled with respect to psychological measures. Such methods have been applied usually to individuals with disorders such as schizophrenia (Curtis et al., 2001; Ebmeier et

al., 1993), and mood disorders, including depression (Bench et al., 1995; Mayberg et al., 1999; Meyer et al., 2001; Videbech et al., 2001). Less attention has been paid to investigating relationships between brain function and personality traits within normal subjects, using these non-invasive methods (Turner et al., 2003).

The non-invasive methods developed, in the last thirty years, to investigate brain function, as distinct from brain structure, include single photon emission computed tomography (SPECT) (Devous, 1989; Prohovnik, 1993), and positron emission tomography (PET) (Ter-Pogossian, 1985). This study in Chapter 5 uses SPECT to measure brain function. SPECT images contain counts of received radiation in each voxel. The count is in proportion to the blood flow at the time of the tracer injection, giving an indication of oxygen uptake, and thus brain function.

While it has been accepted that mental disorders such as schizophrenia, bipolar disorder and depression arise from abnormalities of neurotransmitter function in specific brain regions, it is becoming accepted that 'normal emotional experiences' (Mayberg et al., 1999), and 'normal personality traits' are also related to changes in neurotransmitter function in specific brain regions. For instance, the personality traits of 'detachment', which represents individuals' description of themselves as cold and socially aloof, is related to dopamine D2 receptor density (Breier et al., 1998; Farde, et al., 1997). One now famous model of personality, developed by Cloninger, was explicitly based upon an association of specific personality traits to an underlying neurobiology (Cloninger, 1986; 2008). One of the temperament dimensions in Cloninger's (1986) model, for example, reward dependence, is strongly and negatively correlated to the traits of 'detachment' described above (Breier et al., 1998; Turner et al., 2003).

Image studies help to shine a light into the "black box" (as behaviourists refer to the brain) and they are helping to identify brain areas and systems that underlie important aspects of personality. Ultimately, this information, combined with other kinds of psychopharmacology, neurotransmitter and molecular genetics studies, will lead in time to a detailed understanding of the neurobiological basis of personality. Such an understanding is likely to rival the expectations of earlier, modern personality researchers who studied psychophysiology (Gale, 1973; Strelau, 1983), perception (Witkin et al., 1977), and studied brain lesions (Powell, 1981). As more and more psychologists use revolutionary brain imaging neuroscience techniques, the realization of this vision may be close (Canli et al., 2001; Cloninger, 2002, 2006; Cloninger et al., 1993; Crinion et al., 2007; Cselenyi et al., 2002; Desco et al., 2001; Haier, 2001).

Single photon emission computed tomography (SPECT) has been used to measure brain function, where SPECT images contain counts of received radiation in each pixel or voxel (Devous, 1989;

Matsuda, 2001; Nowak et al., 1986; Prohovnik, 1993). The count is assumed proportional to the blood flow at the time the tracer injection was applied giving an indicator of oxygen uptake, and thus brain function. SPECT brain image data is analysed in Chapter 5.

Personality profiles using normal subjects have been investigated using functional imaging techniques. A recent study (Johnson, 1999) related regional cerebral blood flow (rCBF) to the NEO personality inventory (Costa & McCrae, 1985) and established that the introversion/extroversion personality dimension is significantly related to specific brain regions. Also Turner et al. (2003) showed recently that all seven of Cloninger's personality traits are significantly related to specific areas of the brain.

This study reported here in Chapter 5 uses the technique of statistical parametric mapping (SPM) (Frackowiak et al., 1997, 2003, 2004; Friston et al., 1995a, 1995b; Friston, 2003) to investigate the relationship between regional cerebral blood flow and personality types, in 20 normal males aged between 20 and 33 years, sourced from the study by Turner et al. (2003).

Our work modifies the approach of Turner et al. (2003) by using ICA to denoise the SPECT data, before assessing differences in rCBF between the seven personality traits. ICA decomposes the normalized SPECT datasets into components with maximally independent and spatially fixed regions whose rCBF changes across quartile groups. We will review both the work of Turner et al. (2003) and other literature in personality research and highlight some issues that have progressed (see Section 5.3).

5.2 Materials and Methods

5.2.1 Subjects and Experimental Protocol

This section details the datasets and the protocols used in this study. Recent work by Turner et al. (2003) was the first study to examine the relationships between all seven personality traits of Cloninger's (1994) Temperament and Character Index (TCI) with respect to regional cerebral blood flow (rCBF). Turner et al's study was undertaken in Christchurch New Zealand. This brain function study required ethical approval from the Canterbury Ethics Committee. As a part of the approval process an information sheet was written for potential participants outlining the risks involved with such a study. Ethical approval was given (protocol number 98/05/049 (Turner et al., 2003).

The decision to use male volunteers only was an important aspect of the study design. A small sample size of twenty male volunteers, aged between 20 and 33 years, was recruited from the Christchurch School of Medicine and Health Sciences (CSMHS) and the University of

Canterbury, Christchurch, NZ. The participants were asked to avoid eating or drinking (other than water) in the hours preceding the scan. They were also asked to avoid the use of drugs, such as painkillers, prior to the scan. Volunteers with a history of alcohol or drug problems, head injury, or mental illness were excluded from the study (see Turner, 2004).

Two volunteers were scanned per evening at the Department of Nuclear Medicine, Christchurch Hospital. The volunteers all reported that they were drug free and reported no previous history of alcohol or drug problems, serious head injury or mental illness. Approximately 500Mbq of 99mTc-HMPAO (Ceretec, Amersham Ltd) was injected into the volunteer's blood stream whilst they completed the TCI questionnaire. Volunteers had been instructed to carry on completing the questionnaire while the radiotracer was injected. SPECT imaging was conducted approximately 30-60 minutes post injection to allow for reduction in the background radiation. A dual head rotating gamma camera collected data from 64 views (20 per view) around an 180^0 elliptical arc. A matrix size of 128 × 128 was chosen with a zoom of two. Tomographic reconstruction was completed via filtered back projection with a Butterworth pre-filter, and attenuation and uniformity correction (Turner, 2004). The brain images contain the counts of radiation detected in each voxel. Figure 4.3 in the previous chapter demonstrates the type of information contained in each image. Varying transverse slices have been plotted from slice one, which is the bottom of the brain, up to slice 68, which is the top of the brain (see Figure 4.2). Colour has been used to demonstrate high blood flow and low blood flow areas using the colour bar plotted at the bottom of the Figure 4.3. From this functional image, the general brain structure is visible. The brain images contain more than 500,000 voxels making any analyses difficult and time consuming. Brain images are realigned and normalized into a standard brain space so that comparisons can be made of the same area in each brain across subjects. SPECT brain image data were thus obtained from the work reported in Turner et al. (2003). A brief review of ICA and statistical parametric mapping (SPM) will be given in Sections 5.4 and 5.5.

5.2.2 Personality Scores

Volunteers completed the TCI (Temperament and Character Index) questionnaire (Cloninger, 1994). This psycho-physiological model introduces seven basic traits that measure unique personality axes. These traits are novelty seeking (NS), harm avoidance (HA), reward dependence (RD), persistence (P), self-directedness (SD), cooperativeness (C) and self transcendence (ST). The temperament descriptors (NS, HA, RD, P) are moderately heritable and stable (Cliff, 1987). Character traits (SD, C, ST) on the other hand are described by Cloninger (1994) as referring to the "self concepts and individual differences in goals and values that influence voluntary choices,

intentions and the meaning of what is experienced in life". Therefore, unlike temperament, character changes with age and is influenced by social experiences (Cloninger, 1994; 2003; 2008).

The seven traits are measured by a self-report questionnaire (Cloninger, 1994; 2003; 2005; 2008), the 240 question version of the Temperament and Character Index (TCI) was used in the study by Turner et al. (2003). Each question relates to a single trait. Initial analysis of the questionnaire calculates 25 subscales, which are presented in Turner's thesis (Turner, 2004). Table 5.1 presents the descriptors for high or low scores on each of the temperament traits and the character traits. The descriptions are from Cloninger et al. (1994).

The resulting data from the TCI model were based upon a summation of the scores for each question; in the bivariate form of the self-report which used a zero/one coding system. Calculation of the TCI subscales involved a straightforward mean response for the questions relating to a particular trait. Missing data were accounted for by calculating the mean across only those questions actually answered at an individual level. The levels of the scales within and between individuals were then directly comparable. To calculate the seven TCI traits from the subscales, a mean of the subscales was calculated (Turner et al., 2003; Turner, 2004).

Table 5.1 Descriptors of the temperament traits and the character traits.

Temperament Trait	High Scores	Low scores	Character Trait	High Scores	Low scores
Novelty seeking	exploratory & curious	indifferent	Self- Directedness	mature & strong	immature & fragile
	impulse	reflective		responsible & reliable	blaming & unreliable
	extravagant & enthusiastic	frugal & detached		purposeful	purposelessness
	disorderly	orderly & regimented		resourceful & effective	inert & ineffective
				self-accepted	self-striving
Harm Avoidance	worrying & pessimistic	relaxed & optimistic		habits congruent with long term goals	habits incongruent with long term goals
	fearful & doubtful	bold & confident			
	shy	outgoing	Cooperativeness	socially tolerant	socially intolerant
	fatigable	vigorous		empathic	critical
				helpful	unhelpful
Reward Dependence	sentimental & warm	practical & cold		compassionate & constructive	revengeful & destructive
	dedicated & attached	withdrawn & detached		ethical & principled	opportunistic
	dependent	independent			
			Self -Transcendence	wise & patient	impatient
Persistence	industrious & diligent	inactive & indolent		creative & self-forgetful	unimaginative & self-conscious
	hard-working	gives up easily		united with universe	pride & lack of humility
	ambitious & overachiever	modest & underachiever			
	perseverant & perfectionist	quitting & pragmatist			

5.3 Review of Personality Studies

5.3.1 A review of cerebral blood flow and personality studies

Regional Cerebral blood flow (rCBF) is an important physiological parameter for the examination of brain function. Neuro-physiological measure of brain activity, like cerebral blood flow (CBF) in a resting state, was used in several studies to investigate the biological basis of personality traits (see Table 5.2). Table 5.2 (sourced from Hermes, 2007) gives an overview of these studies. The personality traits most often examined are extraversion/introversion and neuroticism. Some convergence among studies as well as some inconsistencies across studies can be seen in Table 5.2. Personality traits measured are also different across studies in Table 5.2. Not listed in Table 5.2 are the revised Temperament and Character Inventory (TCI–R) by Farmer and Goldberg (2008b). TCI–R is the third stage of development of a widely used multi-scale personality inventory that began with the Tridimensional Personality Questionnaire (TPQ) (Cloninger et al., 1991) and then the Temperament and Character Inventory (TCI) (Cloninger et al., 1994). From Table 5.2 we note specifically the one study of Mathew et al. (1984) that found a relationship between extraversion and introversion, with respect to global baseline CBF in the dorsolateral cortex. This study was based on the largest sample (N=51), but did not analyse the whole brain. Other studies have found associations for specific brain areas but the results are not consistent across studies (Ebmeier et al., 1994; Johnson et al., 1999; O'Gorman et al., 2006; Stenberg et al., 1990; 1993; Sugiura, 2000; Turner et al., 2003).

O'Gorman et al. (2006) found no significant association between baseline rCBF and neuroticism, in agreement with the earlier studies of Ebmeier et al. (1994), Mathew et al. (1984) and Tankard et al. (2003). If one had some insight into the extent that baseline CBF reflects a subject's given trait, and to what degree this association is impacted by situational effects such as phasic changes in the person's hormonal status, or their levels of partial pressure of carbon dioxide, rates of cardiac and respiratory function, or neuronal activity then inconsistencies across studies could be better understood (Ito et al., 2005; Krause et al., 2006). In particular, if baseline CBF only mirrors a trait poorly and is highly variable across different scenarios, then the same subject would demonstrate different CBF measurements across different studies, even though their individual personality traits would be relatively stable (Roberts et al., 2001; Vaidya et al., 2002). The relationship between baseline CBF and personality traits would therefore be dependent on the specific situational conditions that prevail in a given study and some inconsistent results among studies would not be unexpected (Ebmeier et al., 1994; Johnson et al., 1999; O'Gorman et al., 2006; Stenberg et al., 1990; 1993).

Functional neuroimaging is based on modelling relationships between brain activity and personality. Conventional image analysis averages the signal strength measured in nearby voxels, thereby obtaining a reliable measure of the overall level of activity in a given region of the brain. Table 5.2 gives specific studies which investigated the association between personality traits and brain activity regions. We now discuss two brain image studies from Table 5.2.

In a study to investigate the relationship between individual differences in negative affect (NA) and brain activity, Zald et al. (2002) requested subjects who were healthy and were undergoing PET scans, so as to rate the level to which they had felt NA during the month preceding their scans. In two independent samples of subjects, Zald et al. (2002) showed that resting rCBF within the ventromedial prefrontal cortex (VMPFC) was associated with individuals' scores for NA. Moreover, this study demonstrated that variability in basal VMPFC activity across individuals is correlated to individual differences in personal emotional experience.

Tankard et al. (2003) (Table 5.2) demonstrated a significant association between decreased levels of resting dorsolateral blood flow and increased state anxiety responses to a succession of stress provoking stimuli. A significant curvilinear (U-shaped) relationship was also identified between asymmetric dorsolateral perfusion and state anxiety; with higher asymmetric blood flow favouring either the right or the left dorsolateral region - which is recognised to be associated with higher state anxiety levels. This association was conditional on the given subject's age and systolic blood pressure. It was concluded that in older men, in particular, resting perfusion in the

dorsolateral region may be a more accurate and consistent biological marker for state anxiety than trait anxiety (Tankard et al., 2003).

An alternative approach to averaging in functional neuroimaging is to analyse the pattern of variability across voxels in the unsmoothed image (Kriegeskorte & Bandettini, 2007). These capabilities of brain imaging will improve in the future. Current functional neuroimaging techniques like SPECT cannot however determine personality. Brain imaging is only a rough measure of personality, and provides information about traits by analysing activation or deactivation regions in the brain (Haier, 2004; Cloninger, 2002).

5.3.2 The case for nonlinear versus linear rCBF and personality models

Advances in brain imaging technology can be used to help identify the neurobiology and the neuroanatomy of personality traits. In addition this technology is helping to increase our understanding of learning (Haier, 2001; Haier et al., 1992), memory (Alkire et al.,1996; 1998), intelligence (Haier et al., 1988; 2003), and even consciousness (Alkire & Haier, 2001; Alkire et al., 2000). However, whereas brain imaging methods are now a staple of cognitive research, the transition from personality research (based on psychometrics) to research using brain imaging methods, is still rather sporadic and at an early stage.

We now review studies that confirm that the TCI temperament dimensions in fact have non-linear relationships with specific brain regions, not linear as often assumed (see Turner et al., 2003 and Cloninger, 2008). Cloninger (Cloninger, 2005; 2008) recently discussed how factor analysis assumes that variables are linear and how the measured variables are assumed to be on an integral scale, with identical average effects across their range of definition. Invariably this is false for personality variables, as was recently demonstrated in the study of Turner et al. (2003). Turner et al. (2003) clearly showed that linear methods frequently produce inconsistent results in different samples, when they are used to model dynamic systems, where the relationship between the personality traits and regional brain activity is nonlinear. Personality is a complicated representation of nonlinear interactions among an entire hierarchy of learning schemes that have evolved and that develop over time as an adaptive and complex progression - as described in detail by Cloninger (2004) and Kaasinen et al. (2005). Note that evidence of possible nonlinearity and of probable interactive effects was recently given in the study of Gusnard et al. (2003), which demonstrated that whilst the activation of blood flow in a brain circuit involving the ventral striatum, anterior cingulate cortex, and orbitofrontal cortex, in response to increasing frequency of frustrative non-reward, has a positive slope for individuals who score highly on the TCI known as

persistence; the corresponding gradient is negative for subjects with low persistence (Gusnard et al., 2003). Recently Cloninger (2008) has asserted that nonlinearity is in fact a fundamental characteristic of the functional relationships of the entire set of temperament and character dimensions of the TCI instrument, not just of specific dimensions - as was also confirmed by Turner et al. (2003).

Overall the TCI temperament dimensions have strong nonlinear relations with the specific parts of the brain involved in processing negative emotions, habits, and motor skills (Turner et al., 2003). The character dimensions of the TCI also have clear-cut nonlinear relationships with parts of the neocortex which are concerned with the processing of facts and propositions, but not so in relation to the processing of negative emotions and habits (Turner et al., 2003), as has been assumed, to date, in previous research.

Cloninger (2008) acknowledged that summarizing, in brief, the literature on brain function and personality (of over 2,000 peer-reviewed articles, see (https://psychobiology.wustl.edu) is challenging. All seven TCI dimensions have been shown, for example, to have differential associations with different personality disorder subtypes (Goldman et al., 1994; Mulder & Joyce, 1997; Svrakic et al., 1993), with various mental disorders (Cloninger, 1994); and also in relation to distinctive patterns of functional brain activity (Turner et al., 2003) and unique genetic precursors or antecedents (Gillespie et al., 2003).

Indeed strong correlations between the brain function and TCI variables do not guarantee psychological or etiological equivalence. For example, it is dubious that high scores on Harm Avoidance should be combined with low scores on Self-Directedness, as is generally expected , because these qualitatively dissimilar emotional and rational variables have distinctive underlying genetics (Gillespie et al., 2003), neurobiology (Turner et al., 2003), development (Cloninger, 2003), psychometrics, and meaning, as has been discussed above.

The premise that character is correlated with higher cognitive processing in the parts of the brain that are more recently evolved, but not so for temperament (Cloninger, 2002; Turner et al., 2003) is now established . Some studies have demonstrated that the heritability of character is as great as that of temperament (Cloninger, 2004; Gillespie et al., 2003). This significant discovery suggests that the prevailing and popular dichotomy between the neurobiological and psychosocial paradigms of human personality is not warranted (Cloninger, 2008).

Cloninger (2004) and Sugiura (2000) reported significant (linear) correlations between 3 TCI temperament domains and blood flow in several brain regions. However, the study of Turner et al. (2003) provides strong evidence for non-linearity of the TCI-R dimensions. Cloninger (2008) recently acknowledged that Turner et al. (2003) indeed had demonstrated that the TCI temperament dimensions have strong nonlinear relations with the specific parts of the brain involved in processing of negative emotions, habits, and motor skills. In contrast, it was shown that the TCI's character dimensions have strong nonlinear relationships with parts of the neocortex involved in the processing of facts and propositions, not negative emotions and habits (Cloninger, 2008; Turner et al., 2003). In addition, Cloninger (2008) highlighted that Turner et al. (2003) showed that each of the seven TCI dimensions have different relationships with unique patterns of functional brain activity.

Turner et al. (2003) is the only study to date to have examined the relationship between all the seven personality dimensions in the TCI with respect to rCBF. This study is unique in that the imaging occurred while each subject completed the TCI items. Indeed preliminary regression analyses by Turner et al. (2003) did not show a significant linear relationship between rCBF and any of the personality scores. Subsequently a voxel based quartile method was suggested and used, which then revealed many nonlinear relationships, and activations and deactivations between the TCI scales and regional blood flow in a number of areas throughout the brain.

Haier (2004) suggested that because so many relationships between the personality dimensions of the TCI and rCBF were found by Turner et al. (2003), an additional study is necessary with more controllable tasks before Turner's results can be properly evaluated (see also Farmer & Goldberg, 2008a; Zuckerman & Stelmack, 2004). Such an exercise, in fact, is given in this study (Section 5.5), where we adopt a SPM-ICA modelling approach to reanalyse the psychometric brain image (neuroinformatic) data of Turner et al. (2003), to establish if blood flow correlates significantly with scores on all seven TCI constructs for even more regions of the brain.

5.4 Overview of Statistical Parametric Mapping (SPM)

Researchers who are interested in personality have very limited access to imaging technology, which requires collaboration with other specialties, not usually interested in personality research, and requires high budgets, which are virtually unknown in psychometric personality studies. Scanning large samples with tasks appropriate to personality issues would be required along with image analysis techniques that could address multiple problems of anatomical localization and pf statistical inference.

Despite the many difficulties, various methods of brain imaging have identified the functional neuroanatomy of personality traits. As computing technology improved and voxel-by-voxel based analysis became possible, problems grew even more complex and required different and more sophisticated statistical approaches (Zuckerman & Stelmack, 2004). These more modern studies are discussed below and it is noteworthy that the majority of these now use standard image analysis programs like Statistical Parametric Mapping (Friston et al., 1995b; 2003; 2007).

SPM (Friston et al., 2003; 2007), available from the Methodology Group at the Wellcome Department of Cognitive Neurology UK, was developed for just such a statistical analysis of human functional brain imaging data, whether the data was obtained using fMRI, PET, SPECT, EEG or by megneto- encephalogram (MEG) on a voxel -by- voxel basis. SPM provides not only statistical information but also spatial anatomic position for whole brain analysis. Furthermore, spatial normalization is the essential pre-processing step to transform individual images into the MNI (Montreal Neurological Institute, McGill University, California, USA) standard template brain atlas that corresponds with Talairach anatomic coordinates (Nolte, 1981; Svrakic et al., 1992) (see Turner et al., 2003).

SPM is a software toolkit for academics and users who are conversant with the SPM's fundamental statistical, mathematical and image processing ideas to perform analyses of functional imaging data. Effective use of the SPM software depends on such knowledge and for unfamiliar users a recommended text is the introduction to SPM by Friston (2003). More details are also to be found in Friston et al. (2003, 2007) (see http://www.fil.ion.ucl.ac.uk/spm/).

SPMs are so-called image processes with voxel values, which under the null hypothesis, are assumed to be distributed according to a known *a priori* probability density function, traditionally either the Student's T or F distributions. As such these SPMs are referred to as T- or F-maps. SPMs have proved highly successful in applications by virtue of the simplicity of the idea that underlies them, namely, that analysis can be performed at each voxel, one by one, via any traditional, univariate statistical test. The resultant statistical parameters are then compiled into an image - referred to as the SPM. SPMs can be considered to be spatially extended statistical processes, due to the probabilistic behaviour of Gaussian random fields (Adler 1981; Worsley et al., 1992; Friston et al., 1994a), which underpin their formulation. This is by virue of the fact that both the univariate probabilistic features of a SPM and any non-stationary, spatial covariance structure is modelled by Gaussian random fields (GRFs). The rationale is that any anomalous deviations of the SPM may be thought to be regionally specialised effects, which are created by the cognitive or sensorimotor processes that have been affected by the experiment.

The approach was called SPM for three reasons, as given on the SPM web site (see http://www.fil.ion.ucl.ac.uk/spm/). We quote directly (see also Friston et al., 2007):

> "(i) To acknowledge Significance Probability Mapping, and the use of interpolated pseudo-maps of *p* values to summarise the results of multi-channel event-related potential (ERP) studies.
>
> (ii) For consistency with the traditional nomenclature of parametric maps of physiological or physical parameters (*e.g.* regional cerebral blood flow rCBF or volume rCBV parametric maps).
>
> (iii) In reference to the parametric statistics that underpin the maps"

There are three main procedures involved with statistical parametric mapping. The first step is stereotactic normalization and smoothing. This shifts and wraps the brain into a standard brain space so that analysis of a particular voxel is an accurate investigation of the same three-dimensional piece across all the subjects. Smoothing then conditions the data for statistical analysis. The second step is the development and calculation of the GLM which analyses blood flow at each and every voxel, across all the subjects tested, and enables traditional inferences to be made about complex spatially extended data. The GLM is used to estimate parameters that may help explain the variation in spatially continuous data in exactly the manner used for conventional analysis of discrete data. The GLM models the blood flow in each voxel as the dependent variable for a combination of conditions, covariates or nuisance variables. The final step in SPM is modelling at each and every voxel with the aim of finding significant effects, whilst taking into account the high spatial dependence inherent to brain images.

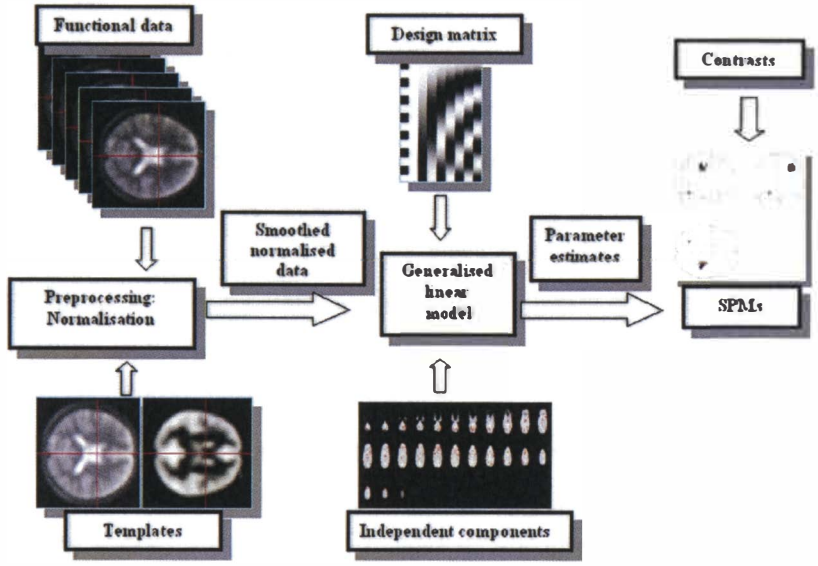


Figure 5.1 Overview of SPM-ICA modelling (http://www.fil.ion.ucl.ac.uk/spm/course/).•

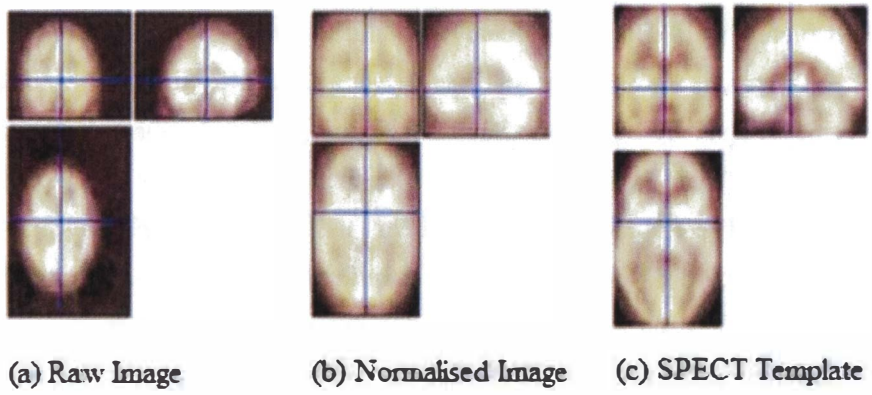


Figure 5.2 Comparison of raw and normalized images versus the SPECT template (Turner, 2004).

Figure 5.1 shows the process steps of our SPM-ICA modelling approach. For more details about the unit of measurement, experimental design, and image pre-processing steps see SPM (http://www.fil.ion.ucl.ac.uk/spm). Further details on ICA for denoising brain images are given in Section 5.5. Details on the quartile based voxel approach used in this thesis, which allows for nonlinear adaptations to the modelling of rCBF with respect to TCI traits, is given in Section 5.6.

Figure 5.2 presents a raw brain image, in addition to the same image after stereotactic normalization and also presents the corresponding template image. Ignoring the scale differences

• The SPM team have kindly allowed us to use materials on their SPM website in this thesis.

of the separate presentations of each image, there is an evident reshaping of the brain to match the template.

Standard brain maps such as the Talairach-Tournoux, or templates from the Montreal Neurological Institute (MNI), are often used to enable research groups worldwide to compare their results. Images are also frequently smoothed (comparable to the 'blur' effect utilised by various image-editing software) in that voxels are averaged with their neighbours, traditionally using a Gaussian filter or also by a wavelet transformation, to denoise the data. By some preliminary cluster results, under different smoothing conditions, 20mm smoothing is chosen since it seems to be the optimum (Turner, 2004).

5.4.1 Statistical Methods of Comparison of Brain Activity

In the SPM framework certain parametric statistical models are assumed at each voxel. In particular the general linear model (GLM) is used to explain variability in the data by including in its formulation terms of both the experimental and confounding effects (in addition to residual variability). Hypotheses of the model parameters are subsequently tested at each voxel via univariate statistics. The SPM approach also enables assessment of the alterations over a given time period (depicting correlations between a task variable and brain activity in a certain brain region) via linear convolution models. These uncover how the signal measured is altered due to changes over time of the subject's underlying neural activity. At a voxel by voxel level, this naturally involves numerous statistical tests, and adjustments need to be done to control for Type I errors (false positives). These are potentially caused by the evaluation of levels of cerebral activity at a very large number of voxels. In typical brain studies, a Type I error would result in falsely detecting background brain activity, and this being interpreted as *activity related to the task*. Adjustments need therefore to be made, based on the number of resolution elements in the image. This is achieved via the theory of continuous random fields, and leads to a description of a new level for statistical significance which corrects for the multiple comparisons problem (Friston et al., 2007).

Divergence in measured brain activity can be depicted in numerous ways. The most basic is a simple table that reports coordinates that display the most significant differences in activity between tasks. Differences in brain activity are conventially shown also as patches of colour on an MRI brain 'slice', with the resulting colours representing the location of voxels that have statistically significant differences between given experimental or intervention scenarios (Friston et al., 2007). Statistical values, such as t-values or z-scores then reflect the gradient of colour.

The relative statistical strength of a given area of activation is in this manner displayed in an insightful and straightforward way (Friston et al, 2007).

5.5 Independent Component Analysis (ICA)

Independent component analysis (ICA) is a computational technique for separating a multivariate signal into additive subcomponents (or hidden factors) under the assumption of mutual statistical independence of the non-Gaussian source signals. It is a special case of blind source separation (BSS) (Hyvärinen, 2005, Lee 1998). The latent variables are called the independent components of the observed data. ICA defines a so-called generative model for the observed multivariate dataset (Stone, 2005; Acharya, 2008).

In the ICA model, the variables are assumed to be linear mixtures of some unknown latent variables, and the mixing system is also assumed unknown. Thus ICA is a technique to separate linearly mixed sources (Hyvärinen et al., 2001b). Superficially ICA is related to both principal component analysis (PCA) and factor analysis (FA). While PCA seeks directions that represent the multivariate dataset best, ICA seeks directions that are maximally independent from each other. ICA is a much more powerful method, however, able to find the underlying factors or sources when these afore-mentioned classic methods (PCA and FA) falter. Data analysed by ICA can derive from many different applications, including finance, digital images, document databases, microarrays, coding studies, astronomy, econometrics, medicine and psychometrics.

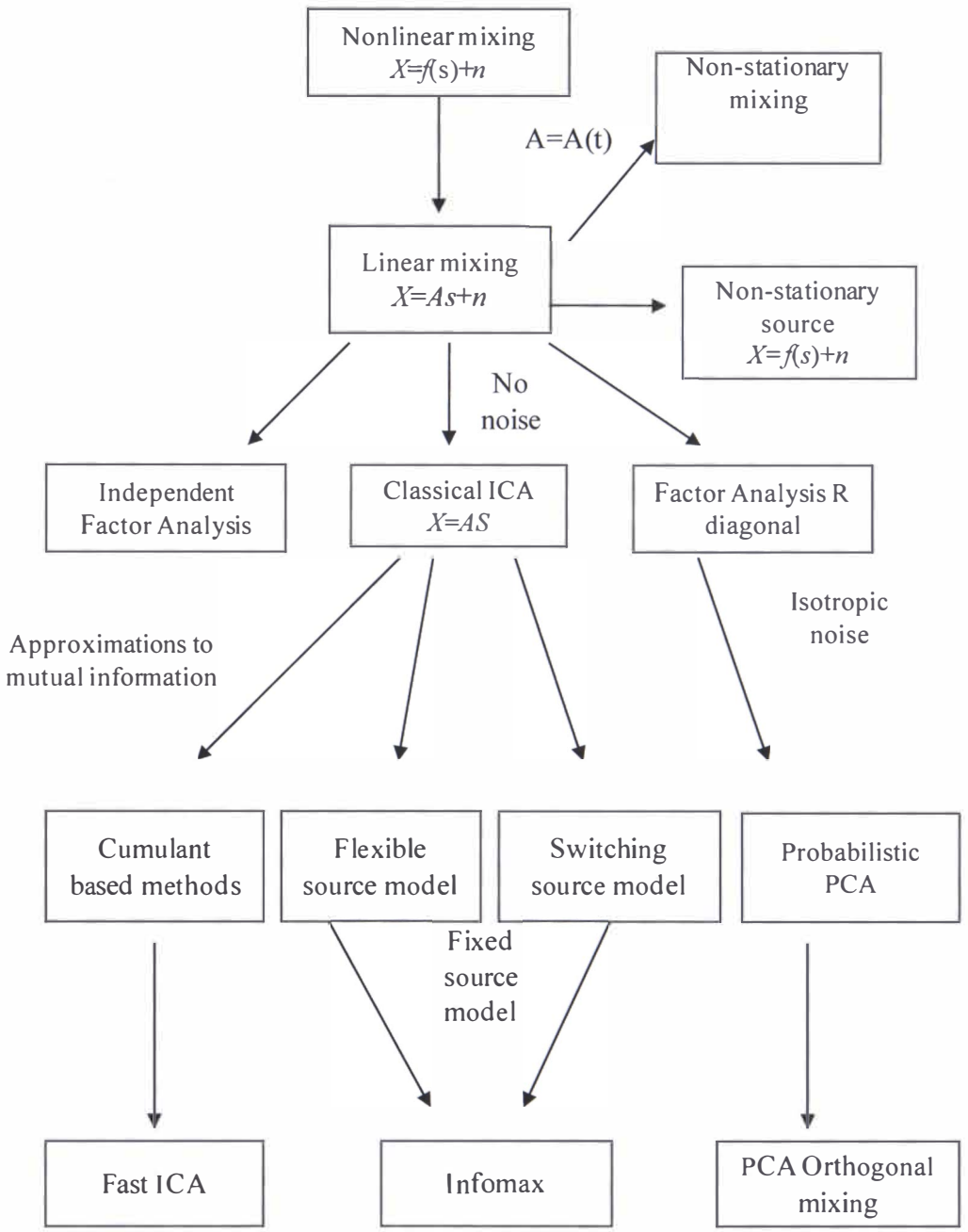


Figure 5.3 Hierarchy of ICA models (Roberts & Everson, 2001).[∞]

Measurements inputted for ICA are often given as a set of parallel signals or time series; in which case the term blind source separation (BSS) (Hyvärinen, 2005) is used to characterise this type of problem. Other typical applications examples are mixtures of simultaneous speech signals (transmitted by several microphones), brain waves recorded by multiple sensors, interfering radio signals inputted into a mobile phone, or parallel time series measured from some industrial

[∞] Professor Roberts (Roberts & Everson, 2001) has given permission to use this diagram in the thesis (personal communication).

process. ICA developed from the blind source separation (BSS) problem, which typically involves recovering original independent sources that have been mixed together. The independent sources $S(t)=\left[s_1(t), s_2(t), \ldots, s_M(t)\right]^T$ are mixed by some matrix, A, to give the observed sources as follows (Roberts & Everson, 2001)

$$X(t) = As(t). \tag{5.1}$$

ICA and BSS aim to find a separating matrix W which recovers the estimated original sources a(t)

$$a(t) = WX(t). \tag{5.2}$$

If W is found such that statistical independence between the components of a(t) is maximised then these are the estimated sources (Roberts & Everson, 2001). For further details about the hierarchy of ICA models, see Hyvärinen et al (2001b), Lee (1998), and Roberts & Everson (2001) (see also Figure 5.4).

Fast ICA (Hyvarinen, 1998) was used for the ICA analysis in this neuroinformatic study. The theory behind FastICA is published in Hyvärinen, (2001a) and Hyvärinen et al. (2001b). FastICA, FA and PCA are part of the broader hierarchy of ICA models (see Page 7 of Roberts and Everson, 2001). A schema depicting this is presented in Figure 5.4 to indicate the relationships between ICA, PCA and FA. Both FastICA and PCA do not allow for a noise term but factor analysis does. PCA has orthogonal mixing, whilst FastICA uses negentropy (normalized differential entropy) minimization to estimate independence.

Further details of the iterative procedure are given in Hyvärinen (2001a), Hyvärinen et al. (2001b), and Roberts and Everson, (2001). A schema is given for convenience in Table 5.3. The FastICA algorithm thus iteratively finds the weight matrix W that maximizes the independence of the estimated sources. This weight matrix represents the loadings of the observed variables onto each component. Independent component analysis forms a family of models, with PCA and FA part of the family (Table 5.3).

An important issue in ICA is the assumption of independence of the underlying latent factors. Shimizu and Kano (2003) studied in some depth the independence of the ICA components. This question is important for the interpretation of components in the psychometric field. Their article initiated the need for the use of test statistics to verify independence of the components and also suggested some alternative statistics that may be useful for establishing independence.

Table 5.2 ICA fixed –point algorithm (Hyvärinen et al., 2001b)

(i) Whiten the data to give $\mathbf{x}$

(ii) Choose m, the number of ICs to estimate. Set counter $p \leftarrow 1$.

(iii) Choose an initial value of unit norm for $\mathbf{w}_p$, e.g. randomly.

(iv) Let $\mathbf{w}_p \leftarrow E\left[\mathbf{x}g\left(\mathbf{w}_p^T\mathbf{x}\right)\right] - E\left[g'\left(\mathbf{w}_p^T\mathbf{x}\right)\right]\mathbf{w}$, where g is defined as given in Equation (5.7).

(v) Do the orthogonalisation $\mathbf{w}_p \leftarrow \mathbf{w}_p - \sum_{j=1}^{p-1}\left(\mathbf{w}_p^T\mathbf{w}_j\right)\mathbf{w}_j$.

(vi) Let $\mathbf{w}_p \leftarrow \mathbf{w}_p / \left\|\mathbf{w}_p\right\|$.

(vii) If $\mathbf{w}_p$ has not converged, go back to step (iv).

(viii) Set $p \leftarrow p$+I. If $p \leq m$, (iii).

5.5.1 Independent Components Analysis of SPECT Data

One application of ICA is feature extraction (Bell & Sejnowski, 1996; 1997; Hoyer & Hyvärinen, 2000; Hurri et al., 1996; 1997; Hyvärinen et al., 1998b; Moulines et al., 1997). Here the columns of A (see Figure 5.4) represent features, and s_i is taken to be the coefficient of the i-th feature in a recorded data vector X (Duann et al., 2003). The theory of redundancy reduction underpins the use of ICA for feature extraction (see Chapter 21 of Hyvärinen et al., 2001b).

In Olshausen & Field (1996), an essentially equivalent method based on sparse coding was applied for the extraction of low-level features of natural image data. The results show that the extracted features correspond closely to those observed in the primary visual cortex (Olshausen & Field, 1996; 1997). These results were seemed to be extremely robust, and have been later replicated by numerous other authors and methods (Bell & Sejnowski, 1997; Hurri et al., 1997; Hyvärinen et al., 1998a; 1998b; Moulines et al., 1997). A systematic comparison between ICA features and the properties of simple cells in the macaque primary visual cortex was conducted by Van Hateren & Van der Schaaf (1998a), and Van Hateren & Ruderman (1998b). These authors found a good match for most of the resultant parameters, especially if video sequences were input into ICA instead of using still images. The features obtained were also closely connected to those found by wavelet theory and Gabor analysis (Comon, 1992; Van Hateren & Ruderman, 1998b). In fact Hyvärinen (1999a) and Hyvärinen et al. (1998b) show how to derive a completely adaptive version of wavelet shrinkage from estimates gained from the noisy ICA model. These features of

data compression and pattern recognition are important and are presented, in part, in this thesis - see Sections 4.3.2 and 5.7.

To analyse SPECT data as shown in Figure 5.5, the rows of the matrix, X, are derived data from multiple individuals (20 subjects in this study) (data sourced from Turner et al., 2003) and the columns are measurements recorded at different voxels (in total 510,340 in this psychoneuroinformatic study). When we compare the SPECT data across a subjects (normals) given in Figure 5.5, we can distinguish differences in rCBF strength, though not that clearly.

Using FastICA, ICA finds an unmixing matrix, $\boldsymbol{W}$, that decomposes the SPECT data into a sum of spatially independent components, $\mathbf{C}=\boldsymbol{W}\mathbf{X}$, and $\boldsymbol{C}$ is a n by v matrix of n spatially-fixed three-dimensional independent component maps. If $\boldsymbol{W}$ is invertible, we can write $\mathbf{X}= \boldsymbol{W}^{-1}\boldsymbol{C}$. The columns of $\boldsymbol{W}^{-1}$ then represent the signal strength of the brain voxels (defined by the component maps), which are used to construct the observed SPECT data, **X**. That is, the signal amplitudes in the columns of $\boldsymbol{W}^{-1}$ represent the relative **rCBF** strength of the brain regions recruited by the corresponding component map across subjects. For example, we expect that some of the resultant components may account for the differences of **rCBF** changes between novelty seeking quartile 1 (NSQ1) and novelty seeking quartile 2 (NSQ2) groups (see Section 5.6.1 for definition of the quartile adaptation of SPM). Other components may account for subject variability in anatomy or in terms of rCBF.

In this study, normalizing each subject's SPECT to the standard brain atlas prior to analysis is very crucial because otherwise ICA will be overwhelmed by subject variability, if normalization is not performed properly. After the ICA training has converged, we apply a simple statistic (the t-test) to the columns of $\boldsymbol{W}^{-1}$ to find components of interest and to test the significance of any difference of **rCBF** between personality groups (quartile groupings) (see Equation (5.17)).

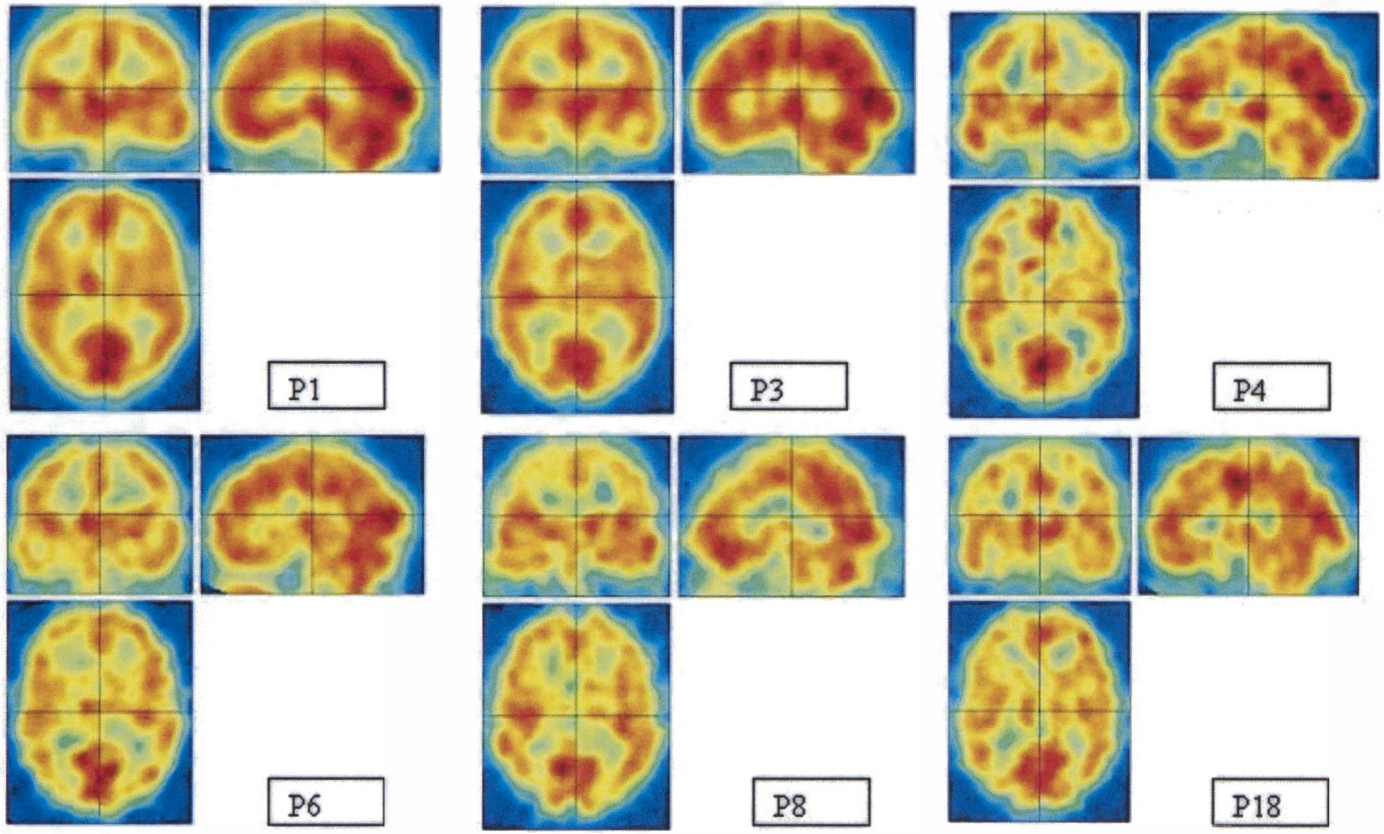


Figure 5.4 Normalized SPECT brain images for selected normal male patients using SPM (Patients 1, 3, 4, 6, 8 and 18 (Pj)).

5.6 SPM –ICA Modelling of Brain Images

As mentioned earlier, statistical parametric mapping (SPM) developed due to the need to analyse functional brain image data that inherently has a large number of responses corresponding to relatively few stimuli. This is a scenario difficult to analyse using standard statistical techniques. The SPM technique used in this thesis, was implemented in SPM5 (www.fil.ion.ucl.ac.uk/spm/software/spm5), which employs the general linear model (Kiebel & Holmes, 2003; Frackowiak et al., 1997; Friston et al., 1995b) at each and every voxel simultaneously, to construct a map characterising the null hypothesis at each voxel. This map is assessed for significant voxels or regions (clusters) (see Section 5.4.1).

After reconstruction, realignment, spatial normalization and smoothing the data is then ready for statistical analysis. This involves two steps: Firstly, statistics indicating evidence against a null hypothesis of no effect at each voxel are computed. An image of these statistics is then produced. Secondly, this statistical image is assessed, so as to reliably locate voxels where a significant effect is exhibited, but at the same time limiting the possibility of false positives. Currently methods for examining data at each voxel are parametric. The parameters are estimated and a given statistic reflecting evidence against the null hypothesis computed. Statistics with a known null distribution are used so that the probability of obtaining significance, given that the null hypothesis is true, can be computed. As such this is hypothesis testing in the conventional parametric sense. Most often the statistical model used is the General Linear Model (GLM) (see Section 5.6.1).

5.6.1 The General Linear Model in Neuroimaging and ICA

The first application of ICA to neuroimaging data was proposed by McKeown et al (1998a; 1998b; 1998c), who demonstrated the usefulness of blind source separation (BSS) for dissecting many types of activity from functional magnetic resonance imaging (fMRI) data. These studies utilised blood oxygen level-dependent (BOLD) contrasts based on an information-maximization algorithm of Bell and Sejnowski (1995; 1996; 1997) to separate and identify regions of the brain related to specific activity.

Conceptually, this algorithm and approach is related to eigenimage analysis through both the execution of both multidimensional scaling and principal coordinate analysis. ICA uses entropy maximization to iteratively uncover spatial modes or their associated dynamics that are approximately independent. The latter condition is a stronger prerequisite than is orthogonality, as

assumed in PCA, and an ICA involves the removal of high order correlations amongst the resultant modes (McKeown et al., 1998a; Friston et al., 2007).

Recall that for our analysis, the rows of the input matrix are SPECT type data and the columns are specific (rCBF) measurements which are recorded at different voxels. ICA finds an 'unmixing matrix, that decomposes or linearly unmixes the SPECT data into a sum of so-called spatially independent components (SICs). We predict that some of the resulting ICA components may explain the differences of the rCBF changes between the seven personality traits groups studied.. Other components (ICs) may also account for subject changeability in anatomy, or in rCBF. In this study, after the ICA training has converged, we apply two sample t-tests to pairs of quartile groups of the personality traits under investigation, and test the significance of the rCBF difference between quartiled personality groupings. Each personality trait thus has 4-quartile groupings (see Equation (5.17)).

Neurobiological dimension models of personality aim to account for the structure of personality traits and disorders through links to neurotransmitter systems (Paris, 2005). As stated before, traditional methods of brain function analysis are based on variants of the linear model (Kiebel & Holmes, 2003; Duann et al., 2003; Friston et al., 1995b; Jung et al., 2001; Papoulis, 1991), as follows;

$$\begin{aligned} X_j &= x_{j1}\beta_1 + x_{j2}\beta_2 + ... + x_{jn}\beta_{1n} + \varepsilon_j \\ &= D\beta + \varepsilon, \end{aligned} \tag{5.13}$$

where

X - $n \times v$ row mean zero data matrix with n the number of patients (subjects)

v – total number of voxels in all slices

D - specified $n \times m$ design matrix, β - $m \times v$ matrix of parameters to be estimated

ε - matrix of noise or residual errors typically assumed to be independent and $N(0,\sigma^2)$.

Once X is specified, standard regression techniques can be used to provide a least squares estimate for the parameters in β. The statistical significance of these can be considered to constitute spatial maps, one for each row in β. These correspond to the specified column of the design matrix.

Using ICA Equations (5.1) and (5.2), we calculate an unmixing matrix, W, to compute spatially independent components, as matrix form

$$a(t) = WX \Rightarrow S = WX, \tag{5.14}$$

where, X is the $n \times v$ row mean-zero SPECT data matrix with n the number of patients (subjects) in the study and v the total number of voxels. W is an $n \times n$ unmixing matrix, and S is an $n \times v$ matrix of n spatially independent component maps.
If W is invertible, we write

$$X(t) = AS(t) \Rightarrow X = W^{-1}S. \tag{5.15}$$

An interpretation of Equation (5.15) is that the columns of W^{1} represent the signal intensity of the brain voxels used to build the observed SPECT data. More explicitly, that the signal amplitudes in the columns of W^{1} signify the relative rCBF strength of the brain regions which are recruited by the corresponding independent component map across individuals (Hyvärinen & Oja, 2000; Lee, 1998; Jung et al., 2000a; 2000b).

The similarity between ICA and GLM can then be seen by comparing Equation (5.13) and Equation (5.14). Starting with Equation (5.14) and performing the initial simple notation substitutions, W into D and S into β, we then have

$$X = D\beta \tag{5.16}$$

which is equivalent to Equation (5.13) but without the Gaussian error term. However there are important differences between Equations (5.13) and (5.16). When the regression Equation (5.13) is used, the design matrix D is specified by the researcher, while in Equation (5.16) the matrix D, computed by the ICA algorithm, also determines β.

Grouping the personality data into quartiles (25^{th}, 50^{th} and 75^{th} percentiles) is an approach that has recently been used in personality studies (Fergusson et al., 2003; Turner et al., 2003). The personality traits grouped (Q1-Q4) are modelled predictors of blood flow in each voxel in this study, and as a GLM of the following form;

$$X_j^k = \beta_{1j}^k x_{1j}^k + \beta_{2j}^k x_{2j}^k + \beta_{3j}^k x_{3j}^k + \beta_{4j}^k x_{4j}^k + \varepsilon_j^k$$

where

$$x_{qj}^k = \delta_{qQ} \qquad \delta_{qQ} = \begin{cases} 1 & q = Q \\ 0 & \text{otherwise} \end{cases} \qquad (5.17)$$

$k = 1,\ldots,K$ voxels in sICs map $j = 1,\ldots,20$ subjects

$q = 1,\ldots,4$ $\beta =$ coefficients

$\varepsilon_j^k =$ error term $Q =$ four quatile groups.

Least squares estimation is used to solve Equation (5.17) to obtain estimates of the coefficients β. Least square estimates minimize the residual sum of squares between Equation (5.13) and Equation (5.17), defined as

$$M_{LSE} = \sum_{j=1}^{J} \left(X_j - x_{j1}\tilde{\beta}_1 - \ldots - x_{j4}\tilde{\beta}_4 \right)^2, \qquad (5.18)$$

where $\tilde{\beta}_1,\ldots \tilde{\beta}_4$ are estimated coefficients and rewritten in matrix notation as follows,

$$D^T X = \left(D^T D \right) \tilde{\beta}. \qquad (5.19)$$

To minimise Equation (5.19) differentiate and set equal to zero as follows

$$\frac{\partial \left(M_{LSE} \right)}{\partial \overline{\beta}_i} = 2\sum_{j=1}^{J} \left(-x_{ji} \right) \left(X_j - x_{j1}\overline{\beta}_1 - \ldots - x_{j4}\overline{\beta}_4 \right) = 0. \qquad (5.20)$$

For a design matrix D that is of full rank ($D^T D$) and invertible, Equation (5.20) can be solved for the least squares estimates of the β coefficients as follows,

$$\tilde{\beta} = \left(D^T D \right)^{-1} D^T X. \qquad (5.21)$$

These $\tilde{\beta}$ are maximum likelihood estimates assuming normally distributed errors.

5.6.2 Contrasts for the General Linear Model (GLM) in SPM

Normally specific brain areas that have increased activity during one specific task or condition versus another need to be established, along with the statistical level of significance of these differences. Contrasts are typically used, as in this study, for such an exercise. I brain imaging studies the form of the signal of interest is not precisely known because the haemodynamic response varies both across individuals and across brain regions. Contrast specification and the interpretation of results are therefore totally contingent on the specification of the model, which also depends on the design of the experiment.

A contrast is a linear combination of the parameters $\mathbf{c}^{T}B$ from the GLM model (Equation (5.13)) $X = D\beta + \varepsilon$ with data X, design matrix D, parameter vector β, and independent errors ε. Recall that SPM creates images whose values are, under the null hypothesis, distributed according to a known probability density function, assumed usually to be the Student's t distribution. These images are known as SPM [T] maps, and are used and described in this study.

Commonly **c** is of the form of a column vector, which delineates a simple contrast of the parameters, being assessed through a so-called SPM [T] in the SPM code. The vector or matrix **c** contains the contrast weights. The null hypothesis equates to the linear combination $\mathbf{c}^{T}\beta$ being zero. The order of the parameters in the parameter (column) vector β, and thus the order to which parameters are referenced in the contrast weights vector **c**, is determined by the formulation of the design matrix (Friston et al., 2003; 2007).

5.6.3 SPM [T] Contrasts in SPM applied to personality trait quartiles

A simple contrast for SPM [T] tests the null hypothesis $\mathbf{c}^{T}\beta=0$ against the one sided alternative $\mathbf{c}^{T}\beta>0$, where **c** is column vector (Poline et al., 2004).

Given an experimental design in which the first two columns of the design matrix relate to the effects of "baseline" and "active" conditions, respectively, then a contrast with weights $\mathbf{c}^{T}$=[-1, +1, 0,...] (and zero weights for any other parameters) tests the null hypothesis that there is no "activation" (i.e. the parameters for both conditions are the same), against the alternative hypothesis that there is some level of activation (i.e. the parameter for the "active" condition is larger than that for the "baseline" condition). The resultant SPM [T] is a statistical image, whose voxel values correspond to the level of the t-statistic for the given contrast at that specific location being modelled. Regions of the SPM [T] with high voxel values therefore provide evidence for an "activation".

Correspondingly, areas of comparative "de-activation" are found by utilising the inverse contrast, namely $\mathbf{c}^T$ =[+1, -1, 0,...]. Likewise, if the third column in the design matrix is a covariate, then the analogous parameter relates to the regression slope; and a contrast with weights c'=[0, 0, 1, 0,...] (with zero weights for all parameters but the third) is then applied.This set up then tests the null hypothesis of a zero regression slope, against the alternative of a positive slope, and is the same as a test of no correlation, against the alternative hypothesis of a positive correlation. This correlation is in reality a partial correlation, delineating the correlation between the data Y and the given covariate, after accounting for all the remaining effects.

In this study we first use SPM to perform an independent two-sample group *t* –test using the quartile groups of the spatially independent components, called sICs, (see Figure 5.6). We then test for significant differences between Q1 and Q2 (Figure 5.6). The contrast is then defined as **c** = [-1, 1] (i.e. Q2 is activated compared to Q1), giving a *t*- contrast test for the *j*th voxel in the sICs map as

$$t_j = \frac{\mathrm{c}\hat{\beta}_j}{\varepsilon_j}, \tag{5.22}$$

where $\varepsilon_j = \sqrt{\hat{\sigma}^2 \mathrm{c}^T \left(X^T X\right)^{-1} \mathrm{c}}$ and the test has [J(total patient) – p(group number)] degrees of freedom, $\hat{\sigma}^2$ is the residual sum of squares divided by the $[J-p]$ degrees of freedom.

Figure 5.6 shows two of the ten ICA components accounting primarily for the rCBF difference attributable to harm avoidance in the SPECT data of Patient 28. The warm colours (red and yellow) indicate increased rCBF, compared with quartile groups of harm avoidance and the cold colours (blue) indicate decreased rCBF in Harm Avoidance.

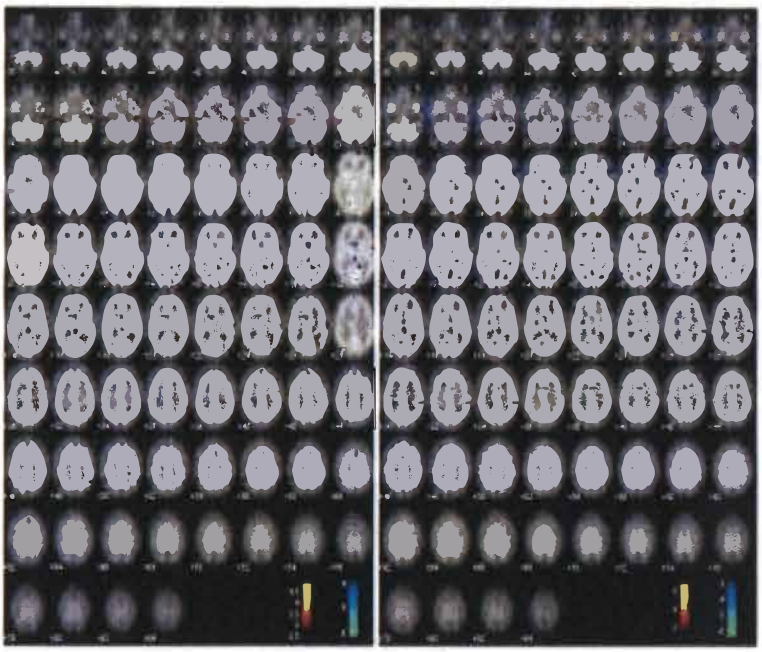


Figure 5.5 Example of two independent components of Harm Avoidance for subject/patient.

Looking at the significance of these tests in every voxel (in this work we assume no covariates) one has to take into account the highly dependent nature of the image data. Seven models were tested, using the seven TCI traits that are grouped into quartiles. These individually are the predictors of regional cerebral blood flow (rCBF). After the t-contrast analysis via the SPM [T]-map, we apply Talairach Daemon (Talairach & Tournoux, 1988), a freely available software on the internet (Lancaster et al., 2000), to convert the coordinates of the brain areas to the anatomical brain atlas.

Every pairwise comparison of Q1, Q2, Q3, and Q4 is tested and only the significant results presented (see Figures 5.9, 5.10, 5.12, 5.14, 5.15, 5.16, 5.17, 5.19, 5.21, 5.22, 5.24, 5.25, 5.26, 5.27, 5.29). A significance level of $\alpha = 0.05$ is used to identify significant clusters. Due to multiple comparisons testing between the four-quartile levels, a Bonferroni correction (Cliff, 1987) is implemented. The Bonferroni correction is generally considered a conservative approach (Brett et al., 2003). Note that spatial correlation is a problem for the Bonferroni correction since this correction assumes that tests are independent. If the voxels are spatially correlated, as in the brain, then the Z scores at each voxel are not independent. This makes the Bonferroni correction conservative (see also Turner (2004)).

5.7 Results: TCI personality and brain function data application

The sample distribution of the TCI personality traits for the 20 normal subjects is presented as side-by-side box plots in Figure 5.7. From Figure 5.7 it can be seen that all the TCI personality traits are non-gaussian (non-normal) and also skewed. All traits are right skewed except HA (Harm Avoidance) and RD (Reward Dependence). The range of the medians across all the temperament traits is within 0.40-0.60, indicating a central tendency for these traits (see also Table 5.4).

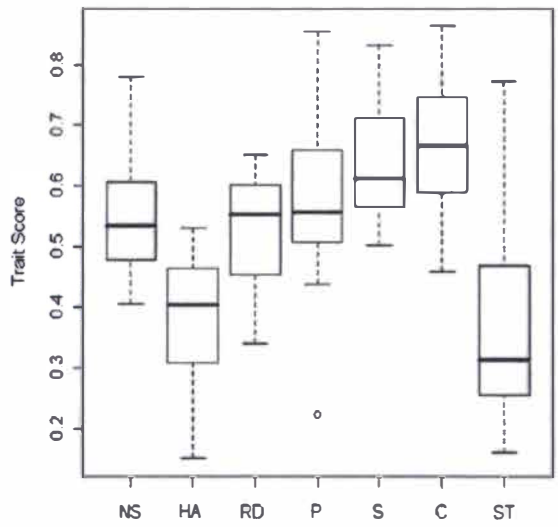


Figure 5.6 Distribution of personality scores for the 20 normal males.

Table 5.4 reports the lower quartile (LQ), median (MD), and upper quartile (UQ) for each personality trait. Significant clusters of activation (a significant relationship between increasing levels of personality scores and increasing blood flow) or deactivation (a significant relationship between decreasing levels of personality scores and increasing levels of blood flow) were found in relation to all seven TCI traits. This is now reported in detail. The results for the SPM_ICA model for NS is detailed in Section 5.6.1, showing the significant clusters found for each TCI trait, grouped by the different quartile contrast.

Table 5.3 Lower quartile, median and upper quartile values for the personality scores.

	Temperament				Character		
	NS	HA	RD	P	S	C	ST
LQ	0.48	0.31	0.45	0.51	0.56	0.59	0.26
MD	0.53	0.40	0.55	0.56	0.61	0.66	0.31
UQ	0.61	0.46	0.6	0.66	0.71	0.75	0.47

The results of the remaining 6 models for HA, RD, P, SD, C and ST are presented in Appendix C, see Tables C.2-C.7. The first column of the relevant contrast Tables gives the contrast (i.e. Q1Q4 which is a contrast between the Q1 and Q4 quartile groups) also specified is, whether the region is an activation or deactivation and below this, the location of the cluster is given in anatomical terms. In regard to Table 5.5 and Tables C.2-C.7, the second column shows the coordinates of the strongest voxel by t-value in the cluster in MNI space (x,y,z), the number of voxels in the cluster(K_E) and finally the p-value associated with observing a cluster of size K_E compared to the expected voxels per cluster (N_E). The bolded p-value are those that are below the Bonferroni corrected significance level ($\alpha \approx 0.008$).

5.7.1 Results for Novelty Seeking (NS)

Novelty seeking seems to be significantly correlated with blood flow, as shown in Table 5.5. Within the novelty seeking model, a contrast between Q1 and Q4 showed a significant relationship with blood flow in the left sub-gyral (K_E = 134). This implies that as the levels of novelty seeking increase from Q1 to Q4, blood flow is significantly increased in the left temporal lobe.

The strength of the regional cerebral blood flow (rCBF) for each NS quartile group is plotted versus NS quartile group in Figure 5.8. This shows general increased blood flow from Q1 to Q4 with NS. Six clusters of activation between quartile groupings Q1 and Q4 were found (Table 5.5). These are now described anatomically.

Cluster K_E = 134 (see Table 5.5) is located in the left sub-gyral showing a significant activation between Q1 and Q4. A cluster of size K_E = 80 is found in the left postcentral gyrus, inferior parietal lobule, which has a significant activation from Q1 to Q4 quartile groups. Cluster K_E = 86 is located in the right sub-gyral and Cluster K_E = 60 is located in the right precentral gyrus. These regions (clusters) all exhibit statistically significant activations. Figure 5.9 (see the three plots on the left hand side) displays the location of the activation clusters between Q1 and Q4, viewed through, the sagittal (top right), coronal (top left), and axial (bottom left) of the brain.

Table 5.4 Results of the contrast analysis for Novelty seeking.

Novelty Seeking	Cluster (x, y, z)	K_E	p-value
Activation Q1Q4			
Left Temporal Lobe: sub-gyral BA20	(-26,-48,16)	134	**0.006**[+]
Left Sub-lobar: lateral ventricle			
Left Parietal Lobe: postcentral gyrus, Inferior parietal lobule,	(-54,-16,22)	80	**0.001**
Right Parietal Lobe: sub-gyral	(26,-32,44)	86	**0.001**
Right Frontal Lobe: Precentral gyrus	(12,-22,64)	60	**0.004**
Right Temporal Lobe: sub-gyral	(40,-32,-4)	63	**0.003**
Left Sub-lobar extra-nuclear	(-24,-32,10)	80	**0.001**
Activation Q2Q4			
Right Frontal-temporal space	(56,16,-4)	473	**0.000**
Left Frontal Lobe: superior temporal gyrus BA22	(-68,12,-4)	110	0.009
Right Frontal Lobe: sub-gyral, middle frontal gyrus,	(22,24,24)	155	**0.003**
Right Frontal Lobe: sub-gyral	(-24,-34,8)	99	0.012
Right Limbic Lobe: anterior cingulate			
Left Frontal Lobe: sub-gyral, middle frontal gyrus	(-26,16,32)	156	**0.002**
Right Temporal Lobe: middle temporal gyrus	(46,-62,4)	106	0.010
Right Sub-lobar: insula BA13	(44,-12,4)	78	0.023
Right Frontal lobe: precental gyrus	(52,4,46)	134	**0.004**
Left Occipital Lobe: lingual gyrus	(-6,-94,-6)	49	0.063
Right Sub-lobar: thalamus, extra-nuclear	(4,-8,2)	112	0.008
Left Occipital Lobe: sub-gyral, cuneus BA18, ligual gyrus	(-28,-68,-2)	189	**0.001**
Left Frontal Lobe: inferior frontal gyrus BA47	(-34,30,-18)	70	0.030
Right Limbic Lobe: cingulate	(8,-26,28)	63	0.038
Left Frontal Lobe: superior frontal gyrus			
Left Parietal Lobe: superior parietal lobule	(-28,-50,62)	48	0.065

[+] Bolded p-values are significant (Bonferroni adjusted)

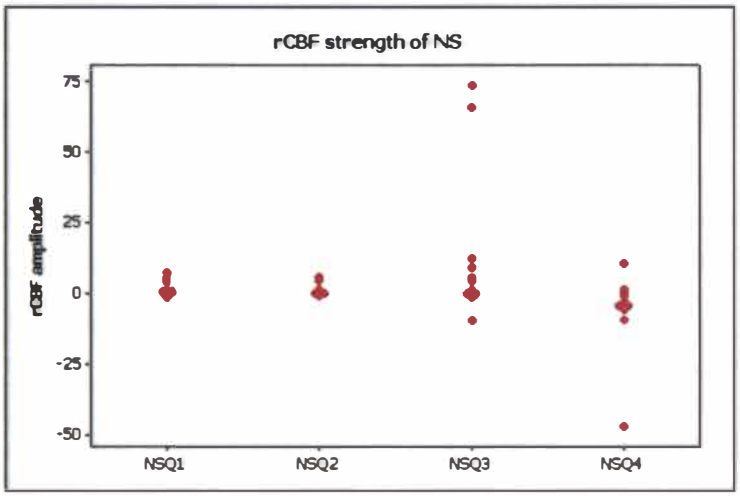


Figure 5.7 Strength of rCBF versus quartiles: Q_j, (j =1, 2, 3, 4) for Novelty Seeking (NS).

Cluster K_E = 473 (see Table 5.5) is associated with a significant increase in rCBF from Q2 to Q4 and is located in the right frontal-temporal space. Graphing the strength of rCBF versus NS quartile group (Figure 5.8) reveals a downward trend from Q2 to Q4. The second cluster K_E = 110 (Table 5.5, Activation Q2Q4) has an associated significant increase in blood flow and this cluster is located in the right superior temporal gyrus (p = 0.003, Table 5.5). Cluster K_E = 155, which overlaps a cluster of activation (K_E = 99) in the right sub-gyral, is also found in the left middle frontal gyrus with a significant activation between Q2 and Q4. Cluster K_E = 99 is also located in the right anterior cingulate (p = 0.012, Table 5.5) (Figure 5.10).

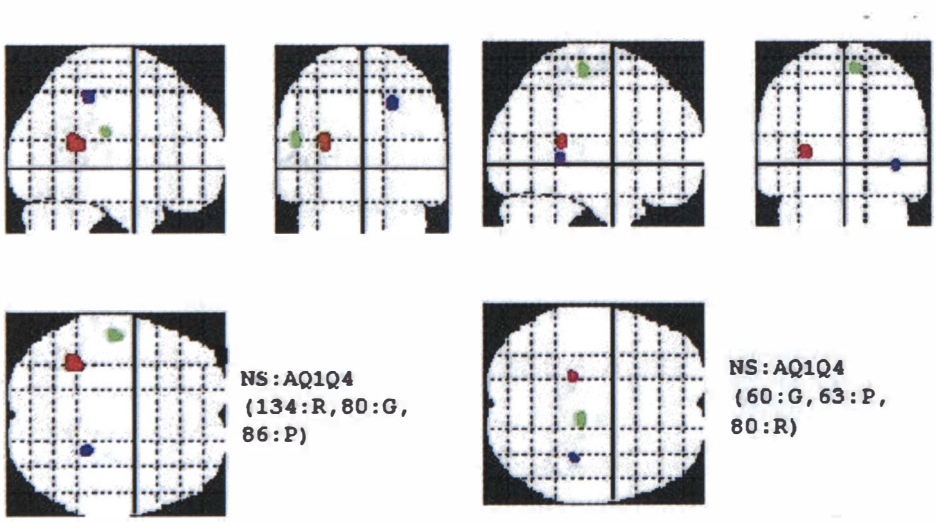


Figure 5.8 Location of activation clusters for novelty seeking (NS) between Q1 and Q4.
(A= activation, R= red, G= green, P= purple).

A cluster of K_E = 156 voxels, found in the left sub-gyral and middle frontal gyrus, is associated with a significant activation from Q2 to Q4 (p = 0.002, Table 5.5). Cluster K_E = 106 is located in the right temporal lobe (middle temporal gyrus) and cluster K_E = 78 is located in the right sub-lobar (insula BA13) showing significant activations (p < 0.03).

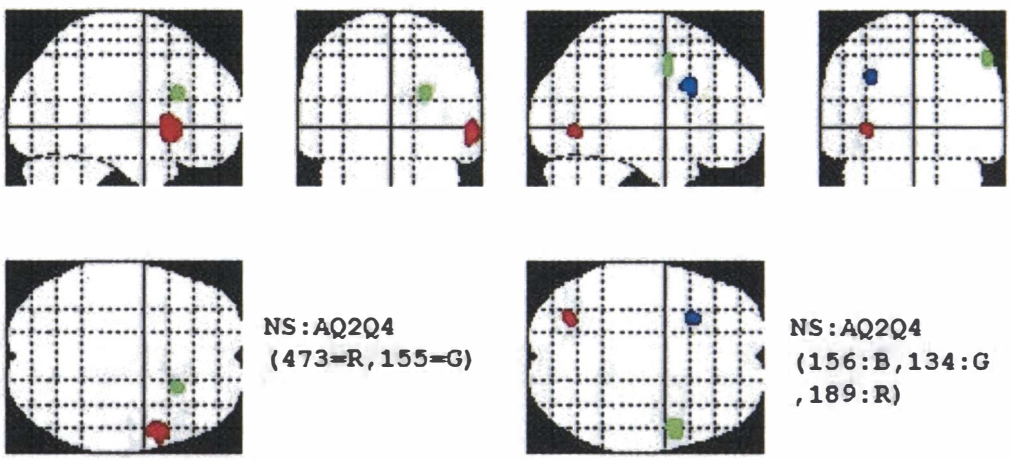


Figure 5.9 Location of activation clusters for novelty seeking (NS) between Q2 and Q4.
(A= activation, R:=red, G=green, B=blue).

Cluster K_E = 189 overlaps with a small cluster of activation (K_E = 49) (p = 0.001, Table 5.5) in the left lingual gyrus, and is associated with a significant increase in blood flow from Q2 to Q4. A small cluster (K_E = 48) of voxels found in the left superior parietal lobule depicts an activation between Q2 and Q4 (p = 0.065, Table 5.5). Figure 5.10 shows the location of these activation clusters between Q2 and Q4 of NS.

5.7.2 Results for Harm Avoidance (HA)

The cluster K_E = 300 is associated with significant increases in blood flow as harm avoidance increases from Q3 to Q4 (see Appendix B, Table B.2). In Figure 5.11, there is a clear increasing relationship between Q3 and Q4. Also small clusters are found between rCBF and harm avoidance quartile groups (see Appendix B, Table B.2) but these are not statistically significant. Figure 5.12 shows the location of the activation clusters between Q3 and Q4 for HA.

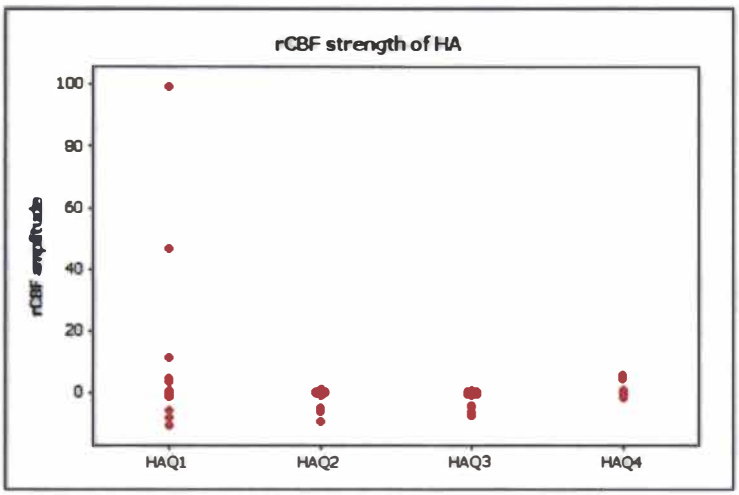


Figure 5.10 Strength of rCBF versus quartile group for Harm Avoidance (HA).

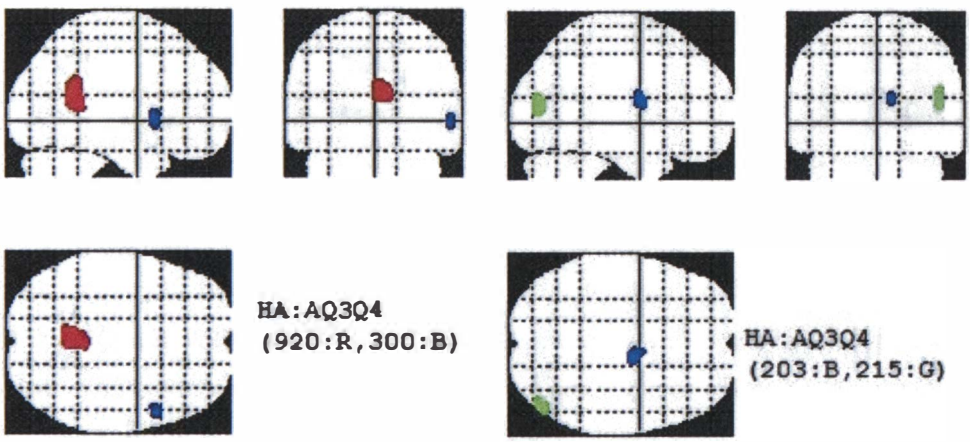


Figure 5.11 Location of activation clusters for harm avoidance (HA) between Q3 and Q4.
(A=activation, R=red, B=blue, G=green).

5.7.3 Results for Reward Dependence (RD)

Table B.3 in Appendix B gives the results of our SPM-ICA modelling of reward dependence (quartiled) as a predictor of blood flow. The cluster $K_E = 316$ has a significant activation from Q1 to Q3. It is located in the right frontal lobe (middle frontal gyrus, sub-gyral and superior frontal gyrus). Figure 5.13 displays the average strength of rCBF versus quartile levels for reward dependence. There is clearly an increasing relationship between Q1 and Q3, but this is not significantly different from Q1 and Q2 (Table C.3). Figure 5.14 shows the location of significant activation clusters between Q1 and Q3 for reward dependence.

Cluster $K_E = 1269$ which is the largest activation region cluster within the reward dependence trait has a significant activation from Q1 to Q3. It is located in the left frontal and limbic lobes (superior frontal gyrus and cingulate gyrus BA32). Cluster $K_E = 298$, which is located in the left

sub-gyral and temporal gyrus BA20, cluster K_E = 321 is located in the right middle frontal gyrus BA6, cluster K_E = 304 in the left sub-gyral and precentral gyrus BA6, and cluster K_E =331 is located in the left postcentral gyrus BA2 and middle frontal gyrus BA9. All show a significant activation from Q1 to Q3. Figure 5.15 shows the location of activation clusters between Q1 and Q4, and between Q2 and Q4 for reward dependence.

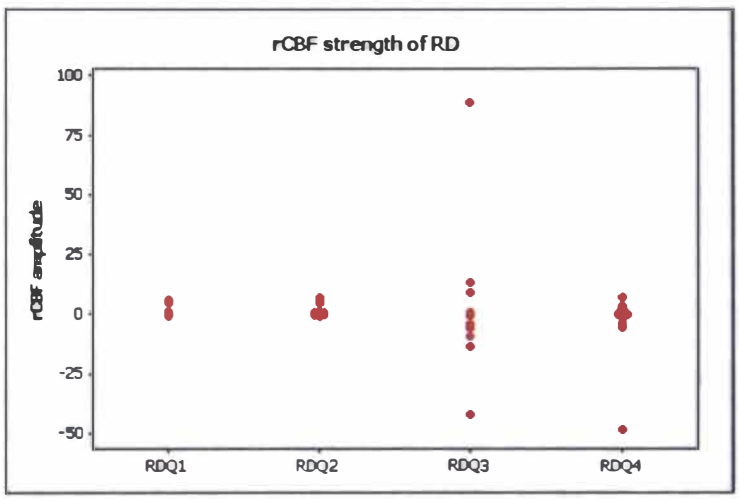


Figure 5.12 Strength of rCBF versus quartile levels for Reward Dependence (RD).

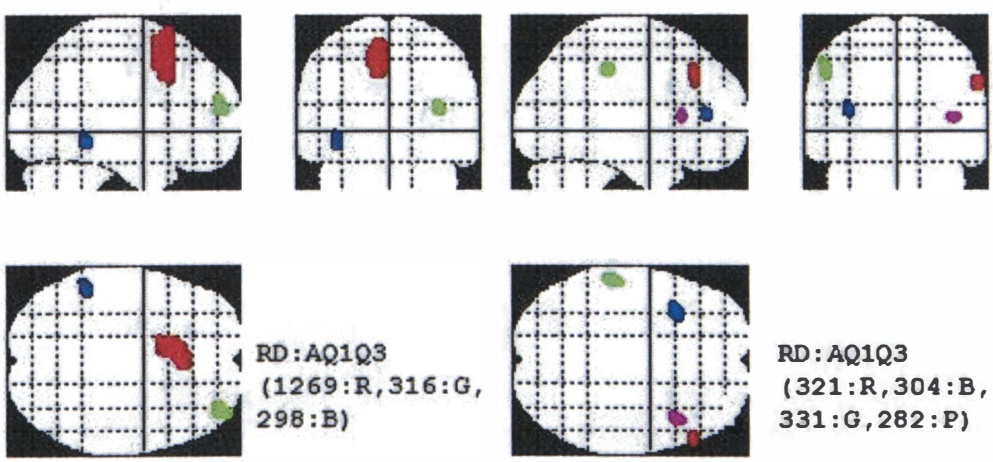


Figure 5.13 Location of activation clusters for reward dependence (RD) between Q1 and Q3.
(A=activation, R=red, B=blue, G=green).

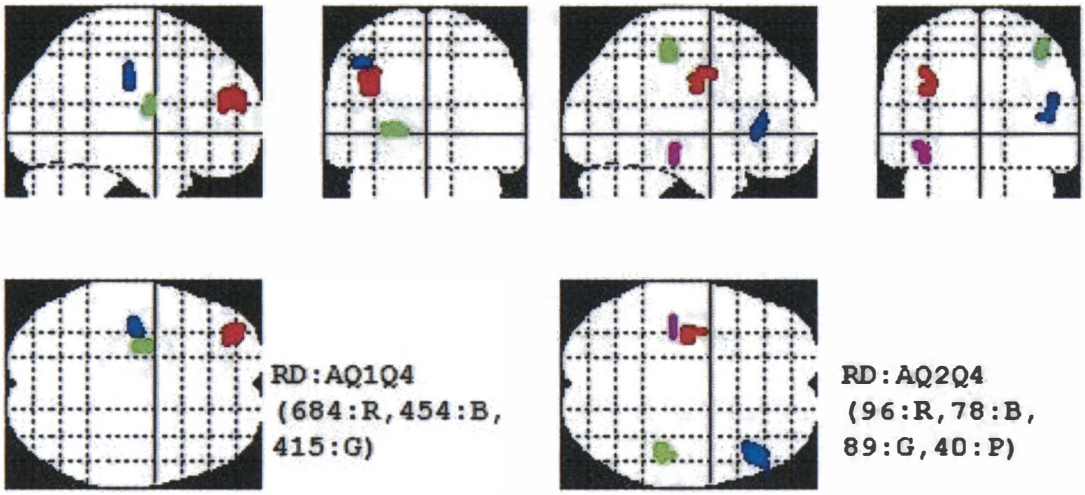


Figure 5.14 Location of activation clusters for reward dependence (RD) between quartile groups.
(A=activation, R=red, B=blue, G=green, P=purple).

Three significant activation clusters are found between Q1 and Q4 for RD. Figure 5.15 shows decreasing blood flow from Q1 to Q4 of RD. Cluster $K_E = 684$ is located in the left middle frontal gyrus BA9, cluster $K_E = 454$ in the left anterior cingulate, and cluster $K_E = 415$ is located in the left lentiform nucleus, extra-nuclear and insula.

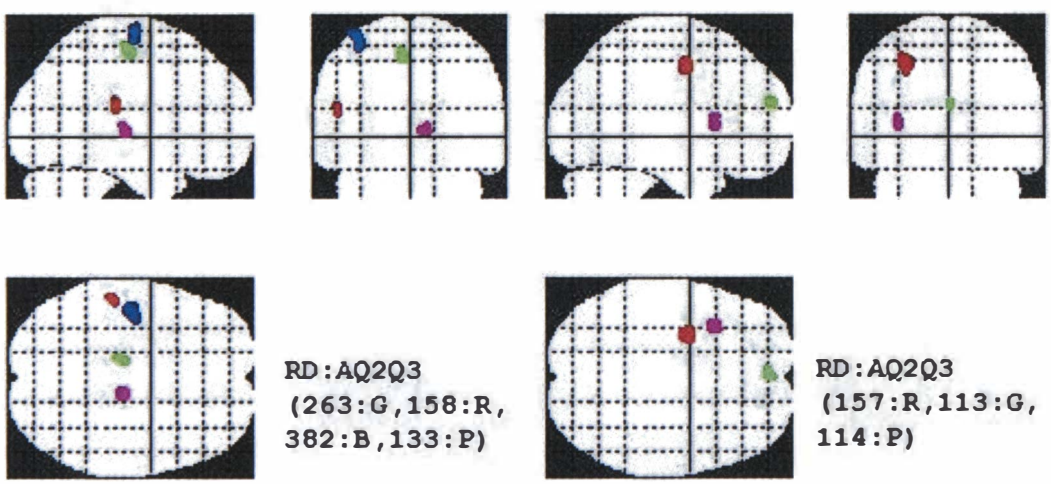


Figure 5.15 Location of activation clusters for reward dependence (RD) between Q2 and Q3.
(A=activation, R=red, B=blue, G=green, P=purple).

Clusters $K_E = 331$ and $K_E = 684$ overlap and in fact $K_E = 331$ is entirely contained in $K_E = 684$. Activations are found in the left frontal (sub-gyral and precentral gyrus) and occipital (lingual gyrus) lobes but these are not significant for Q1 to Q4 (RD) (Figure 5.15).

There are seven significant activations with a large cluster K_E =382 (left insula) between Q2 and Q3. Figure 5.16 indicates that there is a significant increase in blood flow from the second quartile group to the third quartile group for RD. There are three activation clusters: K_E =263 (Left Frontal Lobe: sub-gyral), K_E =158 (Left Parietal Lobe: inferior parietal lobule BA40) and the small cluster K_E =52 (Left Occipital Lobe: middle occipital gyrus), though not statistically significant (see Table C.3 in Appendix C). Figure 5.16 shows the location of activation clusters between Q2 and Q3 for reward dependence.

One deactivation cluster is found between Q1 and Q2 for RD, but this is not statistically significant (shown in Figure 5.17). The relevant cluster K_E =44 is located in the right limbic (parahippocampal gyrus) and temporal (sub-gyral) lobes. Three deactivation clusters are found but these are also not statistically significant from Q2 to Q4 for RD (see Figure 5.17).

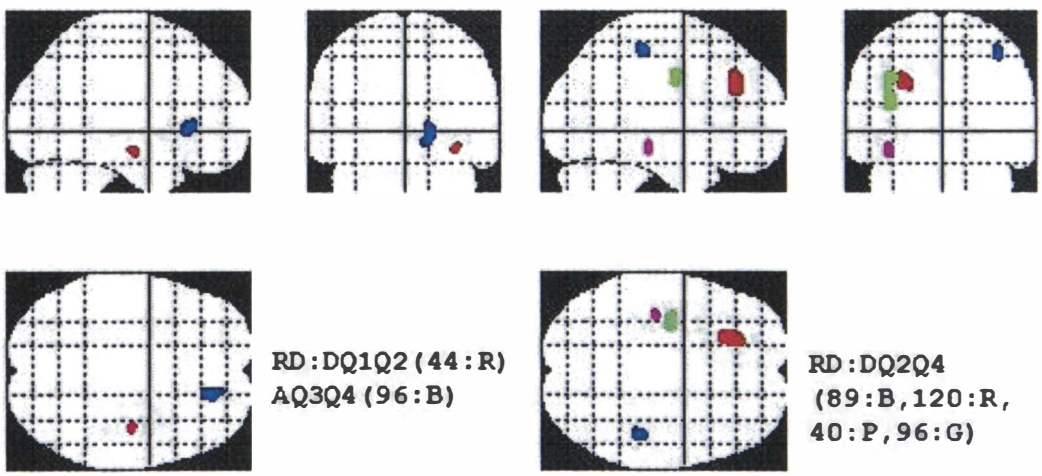


Figure 5.16 Location of deactivation clusters for reward dependence (RD) between Q1 and Q2 and Q2 and Q4. (A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

Cluster K_E =89 is located in the right parietal lobe (postcentral gyrus), and cluster K_E =120 is located in the left frontal lobe (middle frontal gyrus BA8) (see RD: DQ2Q4 (89:B, 120:R, 40:P, 96:G) in Figure 5.17). The smallest cluster K_E =40 is located in the left temporal lobe (Sub-gyral). Figure 5.17 shows also the location of deactivation clusters between Q1 and Q2, and between Q2 and Q4 for reward dependence.

Increasing levels of reward dependence are related to activations in the left frontal lobe (sub-gyral), limbic lobe (parahippocampal gyrus). Also there are activations in the right frontal lobe (inferior frontal gyrus, sub-gyral), parietal lobe (postcentral gyrus) and limbic lobe (anterior

cingulate). These activations are associated with the RD contrast between Q4 and both of Q2 and Q3 (see Figures 5.14-5.17).

5.7.4 Results for Persistence (P)

Table B.4 in Appendix B gives the results of the SPM-ICA modelling of persistence (P) as a (quartiled) predictor of blood flow. Activations are found in the occipital lobe, temporal, frontal lobes, with deactivations found in the temporal, frontal, and limbic, posterior lobes and insula regions of the brain; but these are not statistically significant. Figure 5.18 indicates no significant difference between the four groups in blood flow (with four outliers detected). Figure 5.19 displays the location of the corresponding deactivation and activation clusters for persistence.

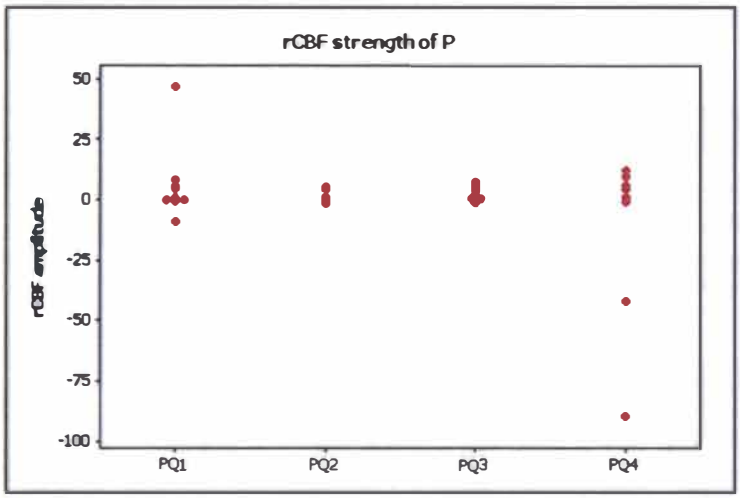


Figure 5.17 Strength of rCBF versus quartiles for Persistence (P).

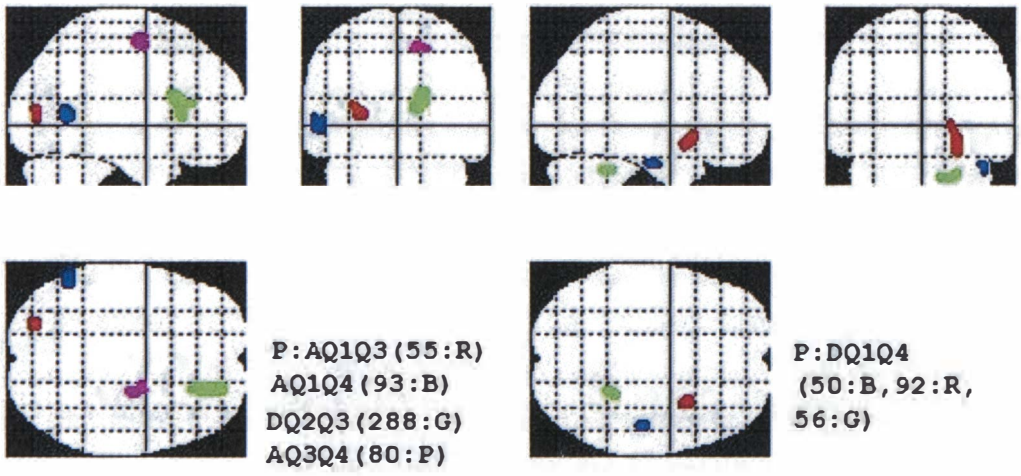


Figure 5.18 Location of activation clusters and deactivation clusters for persistence (P) between quartile groups. (A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

5.7.5 Results for Self Directedness (S)

Table B.5 in Appendix B reports the results of SPM-ICA modelling of (quartiled) self directedness (S) as a predictor of cerebral blood flow. Three activations were found between Q1 and Q3 but were found to be not statistically significant. Figure 5.20 shows the average rCBF strength versus quartile groupings for self directedness.

There is a decreasing relationship of rCBF between Q1 and Q3 for self directedness (Figure 5.20) with a clear decrease of blood flow in the upper quartiles of self directedness. Figure 5.21 gives the location of activation clusters for self directedness. Activations are found in the left anterior (culmen) between Q1 and Q4. Figure 5.20 shows a small decrease in blood flow from Q1 to Q4 for self directedness. The largest activation cluster K_E =1897 is located in the right parietal and temporal lobe regions but is not statistically significant. Cluster K_E =145 located in the left frontal lobe (sub-gyral) and left sub-lobar (insula, extra-nuclear) shows significantly increased rCBF from Q2 and Q3. Figure 5.20 also indicates that there is a significant decrease in rCBF with increasing self directedness from Q2 to Q3. Cluster K_E =118 has a significant activation from Q2 to Q3, and is located in the right frontal lobe (middle frontal gyrus BA9). Cluster K_E = 122, located in the right frontal lobe (middle frontal gyrus BA8, sub-gyral and superior frontal gyrus), and cluster K_E = 362, located in the left frontal lobe (paracentral lobule BA14 and paracentral lobule BA4), have significant activations from Q2 to Q3 (for S). Figure 5.22 shows the location of these activation clusters from Q2 to Q3 for the self directedness trait.

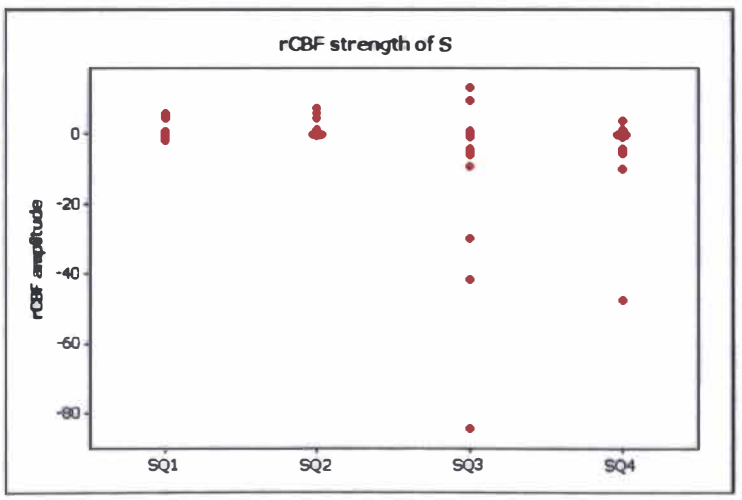


Figure 5.19 Strength of rCBF versus quartiles for Self Directedness (S).

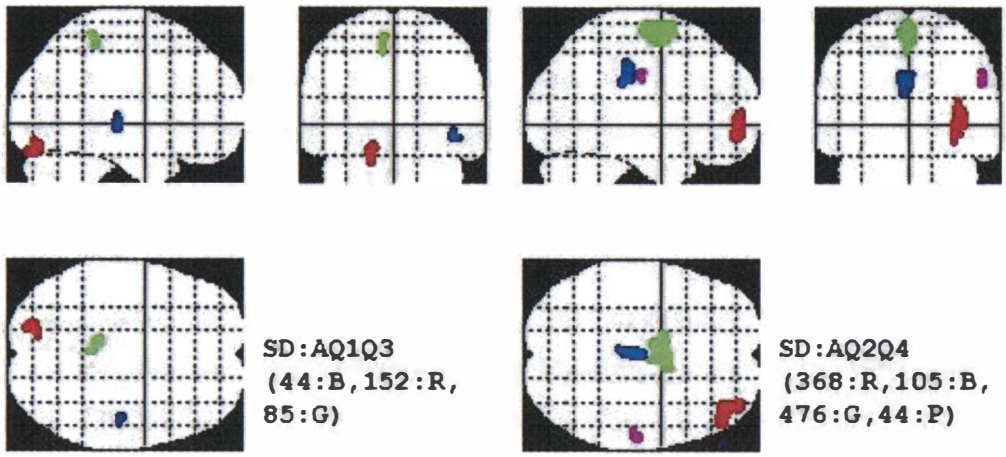


Figure 5.20 Location of activation clusters and deactivation clusters between quartile groups for self-directedness (S or SD).
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

Two clusters (K_E = 368 and K_E = 476) are activations in the right frontal lobe (middle frontal gyrus BA10, sub-gyral) and left frontal lobe (superior frontal gyrus, medial frontal gyrus) between Q2 and Q4 for self-directedness. Figure 5.20 shows little change in rCBF from Q2 to Q4, with no significant difference (and one outlier in Q4). Another two activation clusters (K_E = 105 and K_E = 44) are found between Q2 and Q4 for S, but these are not statistically significant. Deactivations are not found at all, except cluster K_E =40. Figure 5.21 shows the location of activation clusters between Q2 and Q4 for self-directedness, and analogously for the self-directedness Q2 versus Q3 contrast in Figure 5.22.

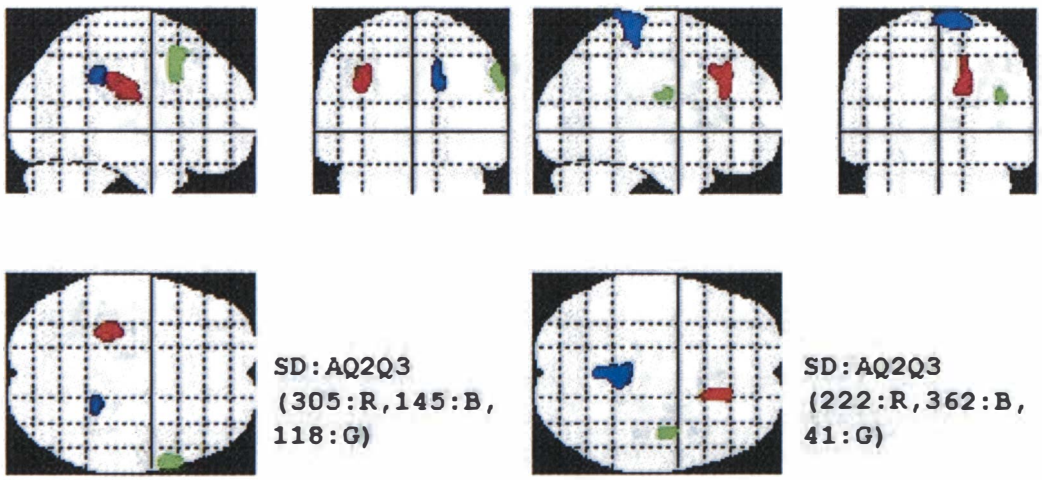


Figure 5.21 Location of activation clusters between Q2 and Q3 for self-directedness (S).
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

5.7.6 Results for Cooperativeness (C)

Table B.6 in Appendix B shows the results of SPM-ICA modelling of the quartiled cooperativeness (C) variable as a predictor of cerebral blood flow. Cooperativeness reveals the most clusters amongst all the seven TCI traits tested. Nine activation clusters are found between Q1 and Q2 (Table C.6). Figure 5.23 indicates there is no difference in blood flow between the cooperativeness quartiles, Q1 and Q2 (except one outlier in Q2). Cluster K_E =96 is located in the left limbic lobe (uncus BA34) and Cluster K_E =41 is located in the left parietal lobe (sub-gyral). The third cluster K_E =82 is a significant activation from Q1 to Q2 (in C) and is located in the right limbic lobe (parahippocampal BA36) (Table C.6). Activation cluster K_E =96 is located in the right occipital lobe (lingual gyrus) and the right limbic lobe (posterior cinulate). Cluster K_E =104 is located in the right occipital lobe (cuneus BA19). Cluster K_E =64 (right limbic lobe: uncus), cluster K_E =96 (left limbic: posterior cingulate, right occipital lobe: ligual gyrus), and the smallest cluster K_E =40 (right limbic lobe: posterior cingulate BA29) are associated with activations between Q1 and Q2 (for C), and all of which are contained in cluster K_E =96 (Table C.6 and Figure 5.24). The location of activation corresponding to all nine clusters between Q1 and Q2 for cooperativeness are shown in Figure 5.24.

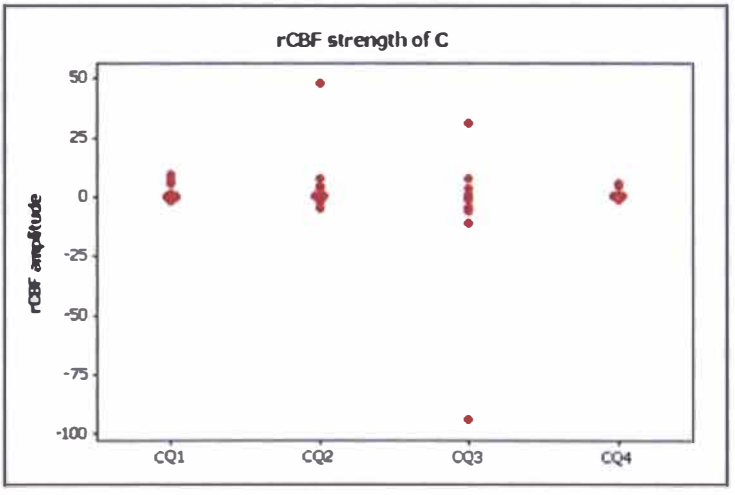


Figure 5.22 Strength of rCBF versus quartiles for Cooperativeness (C).

Cluster K_E =124 (Figure 5.25) is related to a contrast between Q1 and Q4 for cooperativeness, and is located in the left frontal lobe (precentral gyrus). No significant difference is found between Q1 and Q4 in Figure 5.23. Cluster K_E =97 overlaps with cluster K_E =59, which is entirely contained in the larger cluster. In this overlapping region there is a significant activation from Q1 to Q4 for cooperativeness, then a significant activation from Q2 to Q3. Cluster K_E =130 (left anterior lobe: culmen), cluster K_E =64 (left limbic lobe), cluster K_E =54 (left parietal lobe: superior parietal lobule), cluster K_E =108 (left occipital lobe: precuneus), cluster K_E = 45 (left parietal lobe: sub-

gyral), and cluster K_E = 50 (right occipital lobe: middle frontal gyrus and precentral gyrus BA6) all show a significant activation between Q1 and Q4 for cooperativeness (Table B.6 and Figure 5.25).

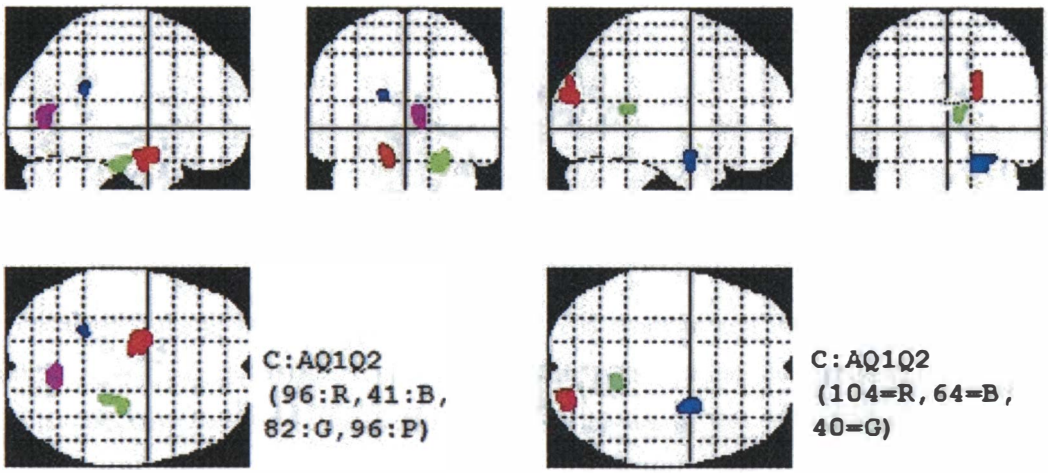


Figure 5.23 Location of activation clusters between Q1 and Q2 for cooperativeness (C)
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple)

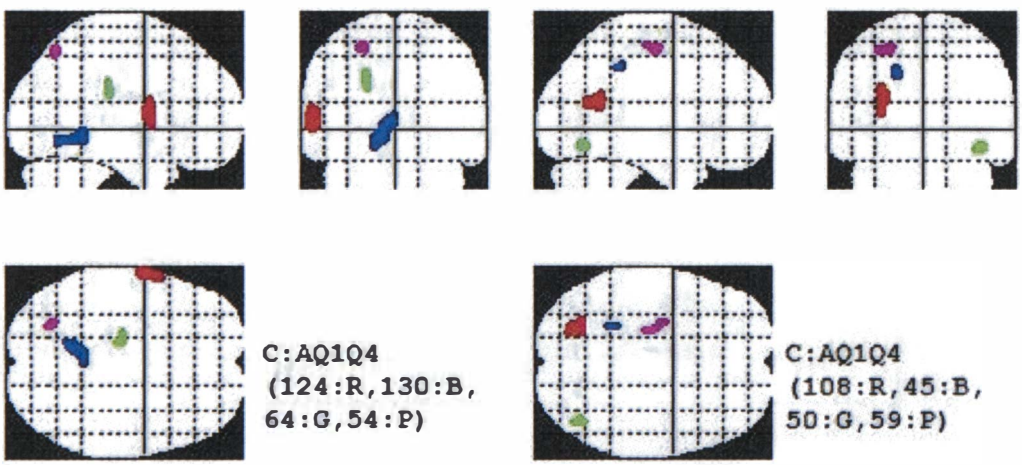


Figure 5.24 Location of activation clusters between Q1 and Q4 for cooperativeness (C).
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

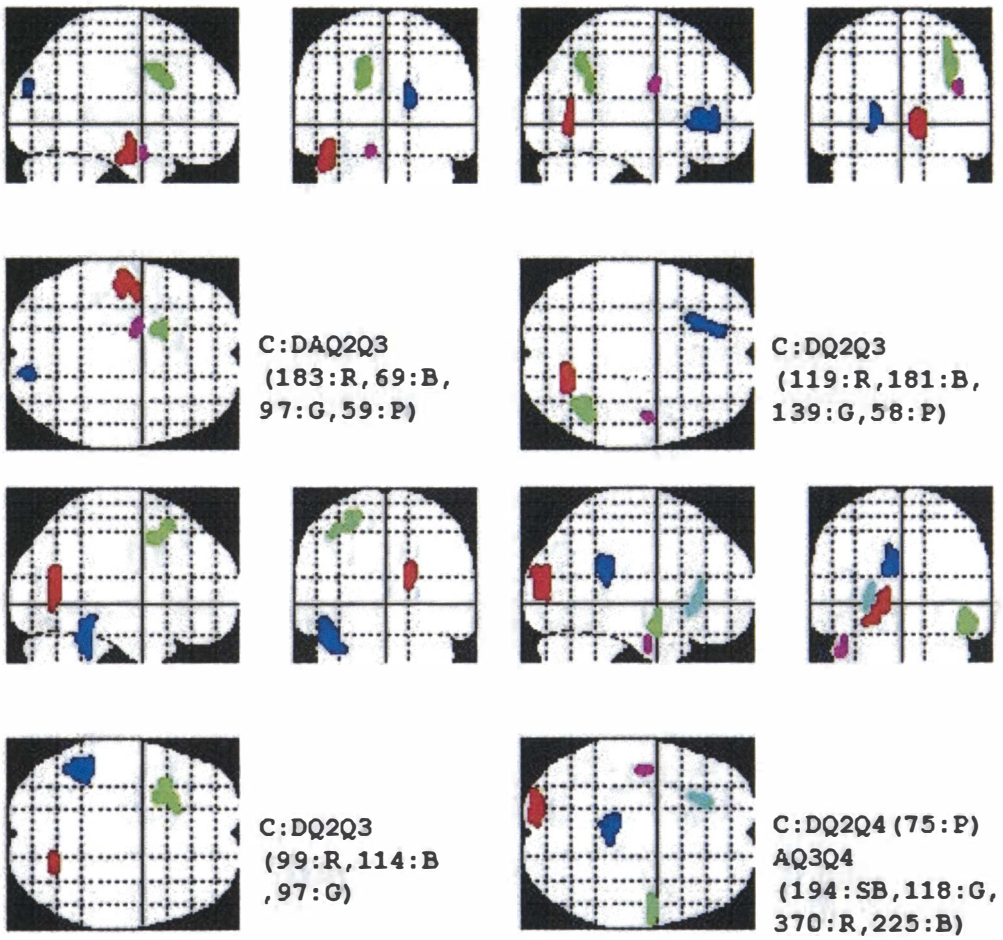


Figure 5.25 Location of deactivation clusters between Q2 and Q3, deactivation between Q2 and Q4, activation between Q3 and Q4 for cooperativeness (C).
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple, SB: sky blue).

Cluster K_E = 183 (Figure 5.26) shows significant deactivations from Q2 to Q3 with respect to cooperativeness and is found in the left temporal lobe (fusiform gyrus and sub-gyral). This cluster overlaps with the cluster of deactivation from Q2 to Q4 (K_E = 75). For these clusters, there is an increased blood flow at low levels of cooperativeness, but increasing cooperativeness is related to decreased blood flow at the high level. Cluster K_E = 181 overlaps with cluster K_E = 69, found in the right occipital lobe (Table C.6). Cluster K_E = 69 is entirely contained in cluster K_E = 181 (Figure 5.26). Both clusters show a significant deactivation between Q2 and Q3 for cooperativeness. Cluster K_E = 97 (left limbic: cingulate gyrus) and K_E = 59 (left limbic: uncus BA34) show a significant deactivation between Q2 and Q3 for C. Cluster K_E = 119 (right occipital lobe: lingual gyrus BA18), cluster K_E = 139 (right parietal lobe: angular gyrus and inferior parietal lobule), cluster K_E = 99 (right occipital lobe: precuneus and cuneus) are all associated with significant deactivations in rCBF between Q2 and Q3 for C.

Cluster K_E = 114 overlaps with K_E = 183, and is found in the left anterior lobe (culmen), left posterior lobe (cerebellar tonsil), and left temporal lobe (fusiform gyrus) with an associated significant activation between Q2 and Q3 for cooperativeness (Figure 5.26). Cluster K_E = 72 (right sub-lobar: extra-nuclear and thalamus), cluster K_E = 58 (right frontal lobe: precentral gyrus), cluster K_E = 55 (right temporal lobe: middle temporal gyrus) and the smallest cluster K_E = 40 (right frontal lobe: sub-gyral) have deactivations from Q2 to Q3, K_E = 43 (right sub-lobar: insula BA13) and cluster K_E = 50 (left parietal lobe: inferior parietal lobule BA40) as well. Figure 5.26 shows the location of deactivation between Q2 and Q3 for cooperativeness.

There is only one cluster K_E = 75 of rCBF deactivation from Q2 to Q4 in the cooperativeness trait. It is located in the left temporal lobe (sub-gyral) which overlaps with K_E = 181. Figure 5.26 gives the location of a deactivation between Q2 and Q4 for cooperativeness.

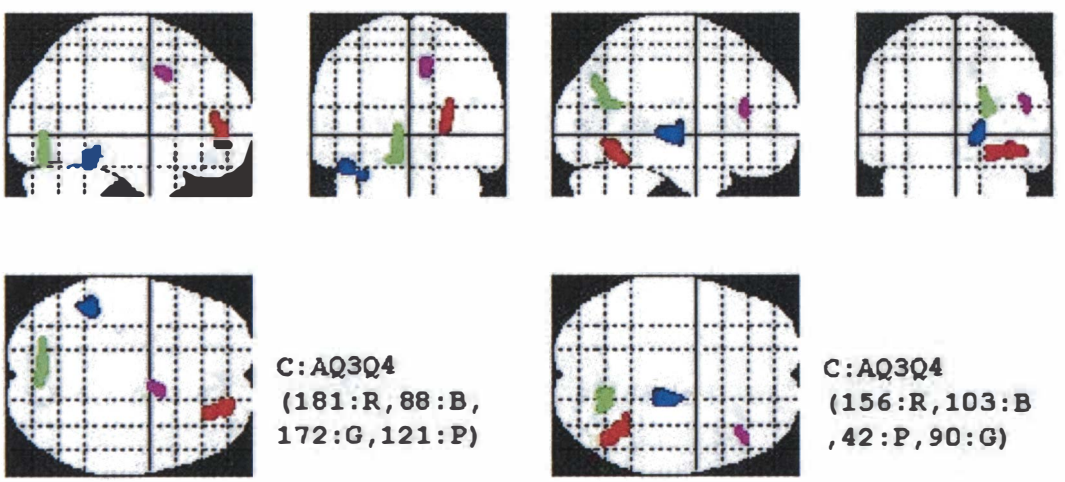


Figure 5.26 Location of activation clusters between Q3 and Q4 for cooperativeness (C).
(A=activation, D= deactivation, R=red, B=blue, G=green, P=purple).

Cluster K_E = 194 is located in the left frontal lobe (sub-gyral) and left limbic lobe (anterior cingulate) showing a significant activation from Q3 to Q4 in C. Cluster K_E = 118 is located in the right middle temporal gyrus BA=21 with a significant activation. The largest cluster K_E = 370 between Q3 and Q4 for cooperativeness is located in the left middle occipital gyrus. This is significantly related to an activation from Q3 to Q4. Cluster K_E = 225 is located in the left cigulate gyrus BA31, left lateral ventile and inter hemispheric and is significantly activated from Q3 to Q4 in the cooperativeness trait. K_E = 181 (right sub-gyral, superior frontal gyrus BA10, and middle frontal gyrus BA9) overlaps with K_E = 382 showing a deactivation between Q1 and Q3, is located in the right sub-gyral, superior frontal gyrus BA10, and middle frontal gyrus BA9, as a significant deactivation. A small cluster K_E = 51, in the right culmen and midbrain shows a significant activation from Q3 to Q4 in cooperativeness.

The mean of rCBF for quartile groups of cooperativeness is plotted in Figure 5.23, which indicates slightly increased blood flow from Q3 to Q4. The next cluster K_E = 88 showed a significant activation between Q3 and Q4 and is located in the left tuber and left inferior temporal gyrus BA37. Cluster K_E = 172 is a significant activation from Q3 and Q4 located in the left declive and inter-hemispheric. Small clusters K_E = 58 (right insula BA47), K_E = 57 (left inferior parietal lobule BA40), K_E = 42 (right lingual gyrus BA 18), and K_E = 46 (right middle frontal gyrus) are found indicating significant activations between Q3 and Q4 in cooperativeness. Cluster K_E = 121 is found in the right cingulate gyrus BA32 and is associated with a significant activation from Q3 and Q4 (in C). Cluster K_E = 156 is also associated with a significant increase in blood flow as cooperativeness changes from 3 to Q4, and is located in the right fusiform gyrus. Cluster K_E = 103 (right thalamus) and cluster K_E = 90 are also associated with significant rCBF activations with the contrast of Q3 and Q4 for cooperativeness.

5.7.7 Results for Self Transcendence (ST)

Table B.7 in Appendix B reports the results for self-transcendence as a predictor of cerebral blood flow. Clusters K_E = 123, K_E = 167, and K_E = 189 are related to the ST contrast between Q2 and Q3 with a significant activation. These three clusters intersect each other. The first cluster is located in the left middle frontal gyrus and the second cluster is located in the middle frontal gyrus and sub-gyral. The third cluster is found in the left sub-gyral. There is a minimal change in blood flow between Q2 and Q3 for ST (Figure 5.28). A large cluster K_E = 801 (Figure 5.29) is located in the left lateral ventricle and inter-hemispheric showing a significant activation from Q2 and Q3. Figure 5.29 displays the anatomical location of the activation regions between Q2 and Q3 for self-transcendence.

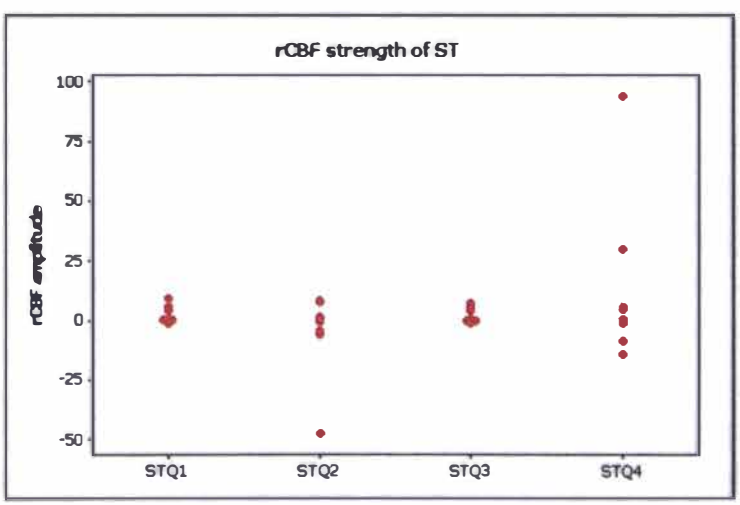


Figure 5.27 Strength of rCBF versus quartiles for self-transcendence (ST).

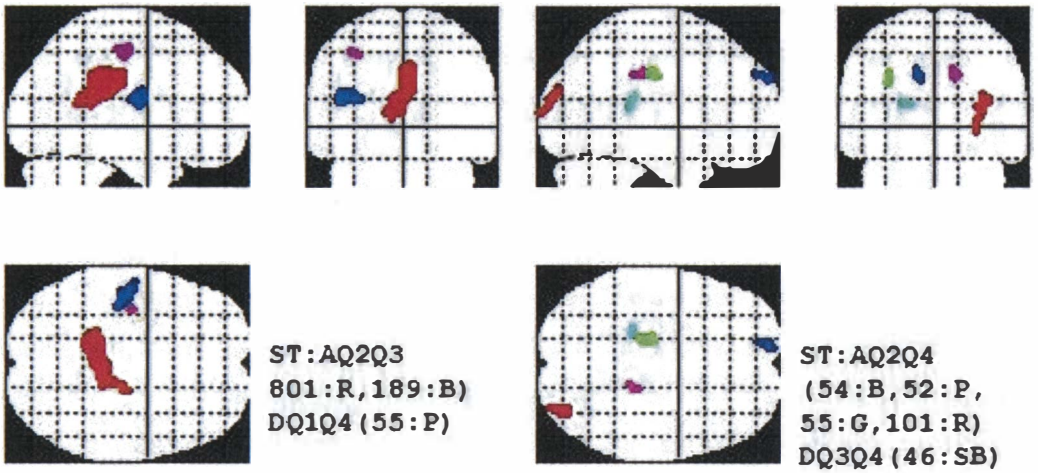


Figure 5.28 Location of activation clusters between Q2 and Q3, and between Q2 and Q4 for self-transcendence (ST). (A=activation, D= deactivation, R=red, B=blue, G=green, P=purple)

There are four small activation clusters between quartiles Q2 and Q4 for self-transcendence. There is a decrease in blood flow from Q2 to Q4 with two outliers (Figure 5.29). Cluster $K_E = 54$ exhibits an activation in the left superior frontal gyrus BA9 and $K_E = 52$ is located in the right cingulate gyrus. A third cluster, associated with a significant activation from Q2 to Q4 is located in the left sub-gyral. The last cluster $K_E = 101$ is found in the right middle occipital gyrus BA 18 and is significant. Figure 5.29 indicates the anatomical location of such activations between Q2 and Q4 for self-transcendence. A small cluster $K_E = 46$ is found in the left lateral ventricle associated with a deactivation from Q3 to Q4.

5.8 Discussion and Conclusions

ICA is shown by this study to be an effective technique to remove artifacts and to separate underlying sources from our SPECT brain signals. In this study, we investigated the relationship between rCBF and all seven of Cloninger's TCI personality traits using ICA followed by SPM. Our results show that a quartile modelling approach with t-contrasts allows for the detection of non-linear relationships between the rCBT and the quartile levels of each of the seven TCI traits. This could not have been detected via linear modelling. This is consistent with the previous work of Turner et al. (2003). Furthermore, this study found more clusters, than the earlier study of Turner et al. (2003), using our SPM-ICA approach.

How is the design of our study valuable? It is noteworthy that Sugiura et al. (2000) used 30 subjects and included both males and females. However they did not test for a gender effect. Our brain psychrometric image study has some major advantages in that only males, were used, so

there was no confounding gender effect. Also the age range in our study (20 to 33 years) is much narrower than that of Sugiura et al. (2000) (26-61 years). Whilst Sugiura et al. (2000) checked for brain atrophy and excluded subjects below a certain threshold, they used however, an eyes closed injection protocol. This causes difficulties in assessing differences due to subject specific thought patterns (at the time the tracer was injected), since what a person thinks about at that point may hugely impact on the blood flow distribution in the brain (see Turner et al., 2003).

Recall that individuals who score highly on novelty seeking (NS) are impulsive, quick-tempered, extravagant and disorderly. In contrast, people who are low in novelty seeking are deemed to be characteristically frugal, rigid, stoical and orderly. Novelty seeking was found to relate to activation in the temporal lobe and right frontal-temporal space, which correspond to the auditory cortex, visual cortex and motor cortex. Our results are thus consistent with previous work by Kolb & Wishaw (1990). They identified eight principal symptoms of temporal lobe damage. These include disturbance of auditory sensation and perception; disturbance of selective attention of auditory and visual input; and disorders of visual perception. Novelty seeking is also expected to be related to the amygdaloid subdivisions on the brain (Wikipedia: http://www.wikipedia.org/), which influence drive-related behaviour patterns and corresponding subjective feelings (Talairach & Tournoux, 1988).

Sugiura et al. (2000) found significant positive correlations between the novelty seeking trait and the orbital prefrontal network. This brain network sends information to the amygdaloid subdivision. Turner et al. (2003) likewise showed an activation related to novelty seeking in the precentral and postcentral gyri, which corresponds to the motor cortex and the somatic sensory cortex. Deactivations were found in the tempotal gyri, occipital lobe and precuneus by Turner et al. (2003). Ito et al. (2005) postulated that increased and decreased brain activations (demonstrated using mutichannel Near-infrared spectroscopy (NIRS)) characterise the cortical reactivities underlying both the novelty seeking and the persistence temperament, respectively. In Ito et al. (2005), oxy-hemoglobin concentration increase was found to be significantly and positively correlated with novelty seeking during the initial time segment of the left-finger tapping task. O'Gorman et al. (2006) found significant associations between novelty seeking and the cerebellum, left thalamus, and cuneus in their voxel-based analysis, which represents a different result to that from to Turner et al. (2003) and also from our results.

Recall that individuals who score high on harm avoidance (HA) are fearful, pessimistic, shy, and fatigable. Those who score low on harm avoidance are generally risk-taking, optimistic, outgoing, and also vigorous. Higher levels of harm avoidance are associated with activations in the limbic lobe and temporal lobe. This is a portion of the brain associated with functions such as olfaction

and emotion. Turner et al. (2003) also found activation in the temporal lobe. Sugiura et al. (2000) demonstrated a negative correlation between harm avoidance and the medial prefrontal network. Two areas of activation were found in the same regions by both Sugiura et al. (2000) and Turner et al. (2003), namely in the right fusiform gyrus and parahippocampal gyrus. However O'Gorman et al. (2006) established harm avoidance has a significant association with the cerebellum, cenenus, and the medial frontal gyrus, regions of the brain.

Characteristically individuals who are high in reward dependence (RD) are approval seeking. In contrast those who exhibit low reward dependence are aloof. Reward dependence is expected to be related to the circuit of Papez (Cloninger, 2001b), which loops from the hippocampal formation, through the fornix, mamillary body, anterior thalamic nucleus, cingulate gyrus, part of the parahippocampal gyrus and back to the hippocampal formation (Nolte, 1981; Turner et al., 2003). This circuit involves the thalamo-cingulate subdivision of the limbic system. Suguira et al. (2000) demonstrated that a significant relationship exists between increased reward dependence and a decrease in cerebral blood flow in the paralimbic regions. Turner et al. (2003) showed that increased reward dependence is related to activations in the frontal and temporal lobes and also similarly with deactivations in the temporal and occipital lobes. Suguira et al. (2000) found an activation in the left middle frontal gyrus, in the same region as a negative correlation. The results presented in this chapter also demonstrateactivations in the frontal, temporal, parietal lobes and sub-lobar for reward dependence. Likewise, deactivations were found in the temporal, parietal and frontal lobes, but were not significant in regard to reward dependence.

Individuals who are high in the persistence (P) traot are known to be determined, perfectionistic, overachivers, as in obsessive personality disorders. In this study, persistence exhibited activations in the occipital, temporal, and frontal lobes, but these activations (clusters) were not statistically significant. Deactivations related to persistence were found in the temporal, frontal, posterior, limbic lobes, and insula. Turner et al. (2003) likewise showed activations in the temporal, parietal, occipital and limbic lobes. Deactivations were found in the parietal, temporal and frontal lobes as well as operculum and insula by Turner et al. (2003). Ito et al. (2005) showed that oxy-hemoglobin concentration increase was significantly and negatively correlated with persistence during the initial time segment of the left-finger tapping task.

The character traits are expected to be related to higher cognitive functions (Cloninger, 2002). Brain regions thus expected to be related to the character traits are the thalamo-neocotical system and the prefrontal cortex (Cloninger, 2002). The prefrontal cortex is the anterior part of the frontal lobes of the brain (http://www.wikipedia.org/). In this study self directedness (S) showed activations in the anterior, frontal lobes, and insula. Deactivations were not found apart from

cluster K_E =40. The anterior pituitary comprises the anterior lobe of the pituitary gland. Under the influence of the hypothalamus, the anterior pituitary produces and secretes several peptide hormones that regulate many physiological processes, including stress, growth, and reproduction (http://www.wikipedia.org/). Turner et al. (2003) also showed activations in the frontal lobe only for self directedness, with deactivations in the precentral gyrus, frontal lobe, temporal lobe and occipital lobe.

We have shown via SPM-ICA modelling that cooperativeness exhibits the majority of clusters (across all seven traits) and that these activations were found in the temporal, limbic, occipital, frontal, and parietal lobes. Deactivations were found in the sub-lobar (extra-nuclear corpus callosum) and parietal lobe. Turner et al. (2003) also showed that most clusters were related to cooperativeness with activations located in the frontal lobe, limbic lobe and sub cortical gray nuclei. Deactivations were found by Turner et al. (2003) to be mainly in the parietal lobe, central region and occipital lobe with involvement also from the temporal lobe, middle frontal gyrus, subcortical gray nuclei, cerebellum and vermis.

We have demonstrated that self transcendence (ST) is related to activations in the parietal lobes and deactivations in the temporal lobe. Turner et al. (2003) showed activations in the occipital lobe and optic radiation and deactivations in the temporal and parietal lobe. The occipital lobe was found to be related to an activation cluster, but was not significant in our SPM-ICA modelling.

This study has presented the results from an investigation of the relationship between brain function and all seven TCI personality traits in twenty normal males using ICA and SPM. Our results show a significant link between regional cerebral blood flow and Cloninger's (2002) personality model, and confirm and extend the results of Turner et al. (2003). Through this study, we have demonstrated a significant relationship between the TCI based psychometric traits and specific regions of the brain. These results support earlier research which reveal a biological basis for the TCI model (Cloninger, 2002; 2008) and for the inherent value of non-linear models in psychometric research, as also recognised recently by Professor Cloninger, founder of the TCI traits (Cloninger, 2008).

Future extensions of this work will be discussed in Chapter 7. This research has the potential to provide constructs for drug trial evaluation of depressives (non-normals) and to help progress the push for patient specific drug design and treatment. Recently Turner, Hudson & Joyce (*In Prep*) have established gender specific and non-linear relationships between the myriad of symptoms of depression and subsets of all seven TCI temperament and character constructs. This work

provides support for our conjecture that the relationships between brain function and TCI may also be gender specific and clearly are non-linear as established by our SPM-ICA modelling, and also as demonstrated by the earlier work of Turner et al. (2003) – the first study, to that date, to investigate all seven of Cloninger's traits. This was also the first study to that date to model non-linear trends between brain function and psychometric constructs. Our work as reported in this chapter finds even more numerous regions of deactivation and activation and also demonstrates clear non-linear dynamic trends.

Chapter 6

6 Detecting Change Points Using Wavelets

6.1 Introduction

Sometimes sudden changes can occur in the time series. Typical change-point problems for a time series $X = X_1,\ldots,X_n$, are concerned with discontinuity for an unknown change-point at X_κ, $(1 \le \kappa \le n)$ if the the observations on the left come from a distribution f and the others come from another distribution f^* (to the right to X_κ).

Many researchers have worked on change-point problems during the past 50 years in various application areas. After a cumulative sum test by Page (1954; 1955) = to detect for a location of change, a change-point problems has been of interest in a variety of areas. Some important papers on detecting change points include; Hinkley (1970), Sen & Srisvastava (1975), Siegmund (1986), they used likelihood ratio approaches, and of Cherno & Zacks (1964), and Smith (1975) used the Bayesian approaches. Carlstein (1988) discussed some nonparametric approaches to the change point problem.

We used wavelets for density estimation in Chapter 3, and wavelet methods are also useful when underlying functions exhibit sharp spikes and jumps that are typical features of interest in the general change point problem. A wavelet approach is based on the idea which links between the function's local homogeneity properties at a certain point and the rate of change of the function's wavelet coefficients near this point across increasing resolution levels (see, for example, Mallat & Hwang (1992), and Ogden (1997)). Local inhomogeneities are identified by increasing or decreasing values of wavelet coefficients at various resolution levels at and around the change point. Wavelet based approaches for change point problems are discussed in more detail in Wang (1995, 1998), Raimondo (1998, 2004) and Park & Kim (2006). Wang (1995, 1998), Raimondo (1998), Cai (1999), Delouille et al. (2001), and Park & Kim (2006) use wavelet estimators for both the detection of change points and the estimation of regression functions. The good time-frequency localization of wavelets provides a natural motivation for their usein change point problems

Specifically this study develops a method to detect a change point in a noisy time series, by performing homogeneity of variance tests on the wavelet coefficients for shifted versions of the times series, to detect the location of change points on a scale-by-scale basis. The shift index is derived (see Section 6.2.4). Significantly, large absolute values of wavelet coefficients indicate that a sudden change in variance has occurred in the noisy time series, indicative of a change point. Our shifted approach known as the shifted (or shifting) SDWT is based on both DWT and MODWT theory. Whilst as detailed in Chapter 6, To motivate the method, wavelet coefficients from simulated data set are examined. The remainder of this Chapter proceeds as follows:

Section 6.2, provides a brief theoretical background on wavelets with some methodological refinements used in the change point detection problem. Section 6.3 presents procedure that is implemented under various scenarios. Section 6.4 compares via simulations our novel procedure with two existing methods: the atheoretical regression trees (ARTs) (Cappelli, Penny, Rea, & Reale, 2007a; Cappelli, Rea, & Reale, 2007b) and the methodology proposed by Bai and Perron (1998, 2003) which is known as the BP. Section 6.5 concludes discussing the advantages of the new proposed method.

6.2 Background and Literature Review

6.2.1 Wavelet Coefficients Near a Change Point

A function has a jump discontinuity at a point t_1 where for a real-valued function $f(t)$ the limit on the left of $f(t)$ as t approaches t_1 and the limit on the right both exist but are significantly distinct.

Consider that $f(t)$ is a function defined on the interval [0, 1], with a jump discontinuity at $t_1 \in (0,1)$ and continuous at all other points in [0, 1]. Here we analyse the behaviour of Haar coefficients when t_1 is inside or outside of the dyadic interval $I_{j,k}$, which is defined below. In particular, we can find the location of a jump discontinuity just by examining the absolute value of the Haar coefficients.

Assume that $f(t)$ is C^2 if $f(t)$ is two-times continuously differentiable on $[0,t_1]$ and $[t_1,1]$. This means that both f' (first differentiation) and f'' (second differentiation) exist, are continuous functions, and hence are bounded on each of these intervals. For fixed $j \geq 0$ and $0 \leq k \leq 2^j - 1$, let $t_{j,k}$ be the mid-point of the interval $I_{j,k}$ that is $t_{j,k} = \left(k + \frac{1}{2}\right)2^{-j}$ since

$I_{j,k} = \left[k2^{-j}, (k+1)2^{-j} \right)$ which is known as a dyadic interval. Two cases are considered for t_1, namely, when $t_1 \notin I_{j,k}$ (t_1 outside the dyadic interval) or $t_1 \in I_{j,k}$ (inside dyadic interval). We also define $I_{j,k} = I'_{j,k} \cup I''_{j,k}$, where $I'_{j,k}$ and $I''_{j,k}$ are dyadic intervals at scale $j+1$, to denote the left half (LH) and right half (RH) of the interval $I_{j,k}$. In fact, $I'_{j,k} = I_{j+1,2k}$ (LH) and $I''_{j,k} = I_{j+1,2k+1}$ (RH).

Now we examine the first case for $t_1 \notin I_{j,k}$, then expanding $f(t)$ about $t_{j,k}$ by Taylor's formula (Vidakovic, 1999; Walnut, 2004; Nason, 2008);

$$f(t) = f(t_{j,k}) + f'(t_{j,k})(t - t_{j,k}) + \frac{1}{2} f''(\mu_{j,k})(t - t_{j,k})^2, \qquad (6.1)$$

where $\mu_{j,k} \in I_{j,k}$. Now using the wavelet property $\int \psi_{j,k}(t)\, dt = 0$ and that $f(t)$ is defined as continuously differentiable then we have the inner product

$$\begin{aligned} \langle f, \psi_{j,k} \rangle &= \int_{I_{j,k}} f(t) \psi_{j,k}(t)\, dt \\ &= f(t_{j,k}) \int_{I_{j,k}} \psi_{j,k}(t)\, dt + f'(t_{j,k}) \int_{I_{j,k}} \psi_{j,k}(t)(t - t_{j,k})\, dt \\ &+ \frac{1}{2} \int_{I_{j,k}} f''(\mu_{j,k})(t - t_{j,k})^2 \psi_{j,k}(t)\, dt \\ &= f'(t_{j,k}) \int_{I_{j,k}} t\ \psi_{j,k}(t)\, dt + r_{j,k}(t). \end{aligned} \qquad (6.2)$$

If we take an absolute for the remainder $r_{j,k}(t)$ of Equation (6.2) to approximate, we then have

$$|r_{j,k}(t)| = \frac{1}{2} \left| \int_{I_{j,k}} f''(\mu_{j,k})(t - t_{j,k})^2 \psi_{j,k}(t)\, dt \right|.$$

Solve the middle part of Equation (6.2) as below

$$\int_{I_{j,k}} t\ \psi_{j,k}(t)\,dt = \int_{k2^{-j}}^{(k+1/2)2^{-j}} 2^{j/2} t\ dt - \int_{(k+1/2)2^{-j}}^{(k+1)2^{-j}} 2^{j/2} t\ dt$$

$$= 2^{j/2}\left\{\left[\frac{t^2}{2}\right]_{k2^{-j}}^{(k+1/2)2^{-j}} - \left[\frac{t^2}{2}\right]_{(k+1/2)2^{-j}}^{(k+1)2^{-j}}\right\} \qquad (6.3)$$

$$= -\frac{1}{4}2^{-3j/2}.$$

From Equation (6.2) and (6.3), we obtain $\langle f, \psi_{j,k}\rangle = -\frac{1}{4}2^{-3j/2} f'(t_{j,k}) + r_{j,k}(t)$.

By using Cauchy's estimate, then the absolute remainder $r_{j,k}(t)$ (Vidakovic, 1999; Walnut, 2004), namely $|r_{j,k}(t)|$, will be

$$|r_{j,k}(t)| \le \frac{1}{2}\max_{t\in I_{j,k}}|f''(t)| \int_{I_{j,k}} (t - t_{j,k})^2 |\psi_{j,k}|\,dt$$

$$\le \frac{2^{j/2}}{2}\max_{t\in I_{j,k}}|f''(t)| \int_{k2^{-j}}^{(k+1)2^{-j}} (t - t_{j,k})^2\,dt$$

$$= \frac{1}{24}.2^{-\frac{5j}{2}}.\max_{t\in I_{j,k}}|f''(t)|. \qquad (6.4)$$

For large value of j, $2^{-\frac{5j}{2}}$ in Equation (6.4) is very small compared with $2^{-\frac{3j}{2}}$ in Equation (6.3). Then we can approximate

$$|\langle f, \psi_{j,k}\rangle| \approx \frac{1}{4}.2^{-3j/2}|f'(t_{j,k})| = O\left(2^{-\frac{3j}{2}}\right). \qquad (6.5)$$

Now we consider the second case, t_1 inside the dyadic interval $t_1 \in I_{j,k}$. If $t_1 \in I_{j,k}$, then either it is in $I_{j,k}^m$ or in $I_{j,k}^n$. We assume that $t_1 \in I_{j,k}^m$, and the other case is similar. Expanding $f(t)$ about t_1 by Taylor's formula, we have

$$f(t) = f(t_1^-) + f'(\mu^-)(t - t_1),\ t \in [0, t_1) \text{ and } \mu^- \in [t, t_1]$$
$$f(t) = f(t_1^+) + f'(\mu^+)(t - t_1),\ t \in [t_1, 1) \text{ and } \mu^+ \in [t_1, t].$$

The Haar coefficient can then be calculate as follows:

$$\begin{aligned}\langle f, \psi_{j,k} \rangle &= \int_{I_{j,k}} f(t)\psi_{j,k}(t)\,dt \\ &= 2^{j/2} \int_{k2^{-j}}^{t_1} f(t_1^-)\,dt + 2^{j/2} \int_{t_1}^{(k+1/2)2^{-j}} f(t_1^+)\,dt - 2^{j/2} \int_{(k+1/2)2^{-j}}^{(k+1)2^{-j}} f(t_1^+)\,dt + \varepsilon_{j,k} \quad (6.6) \\ &= 2^{j/2}(t_1 - k2^{-j})\{f(t_1^-) - f(t_1^+)\} + \varepsilon_{j,k}.\end{aligned}$$

The absolute reminder (Vidakovic, 1999; Walnut, 2004) is

$$\begin{aligned}|\varepsilon_{j,k}| &\le \max_{t \in I_{j,k} \setminus \{t_1\}} |f'(t)| \int_{I_{j,k}} |t - t_1| |\psi_{j,k}(t)|\,dt \\ &\le 2^{j/2} \max_{t \in I_{j,k} \setminus \{t_1\}} |f'(t)| \int_{I_{j,k}} |t - t_1|\,dt \qquad (6.7) \\ &= \frac{1}{4} \max_{t \in I_{j,k} \setminus \{t_1\}} |f'(t)|\ 2^{-3j/2}.\end{aligned}$$

For large j, $2^{-3j/2}$ in Equation (6.7) is very small compared with $2^{-j/2}$. Therefore we can approximate

$$\langle f, \psi_{j,k} \rangle \approx 2^{j/2} |t_1 - k2^{-j}| |f(t_1^-) - f(t_1^+)|. \qquad (6.8)$$

The value of $|t_1 - k2^{-j}|$ is very small if t_1 is close to the left end point of $I_{j,k}^m$ and can be set to zero. However, if we consider the case where t_1 will be in the middle of $I_{j,k}^m$ so that $|t_1 - k2^{-j}| \approx \frac{1}{4}.2^{-j}$. Thus, for large j,

$$|\langle f, \psi_{j,k} \rangle| \approx \frac{1}{4}.2^{-j/2} |f(t_0^-) - f(t_0^+)| = O\left(2^{-\frac{j}{2}}\right). \qquad (6.9)$$

Comparing the decay order of $O\left(2^{-\frac{3j}{2}}\right)$ with $O\left(2^{-\frac{j}{2}}\right)$, we see that the decay of $\left|\left\langle f,\psi_{j,k}\right\rangle\right|$ for large j is considerably slower if $t_1 \in I_{j,k}$ than if $t_1 \notin I_{j,k}$. That, it is the large coefficients in the Haar expansion of a function f that suggest the presence of a jump discontinuity in the interval $I_{j,k}$ correspond also to large coefficients for all scales where $j \geq 0$ and $0 \leq k \leq 2^j - 1$. We conclude that large wavelet coefficients inform us there is a significant jump at a change point in the series.

For other processes, we briefly introduce the discrete Haar transform (DHT) which entails changing integrals in Equations (6.1)-(6.9) into matrix form see more details Ogden (1997), Vidakovic (1999), Percival & Walden (2000), Walnut (2004), and Nason, (2008). The DHT gives strong motivation for our use of the DWT in the detection of a change point.

6.2.2 The relationship between wavelet scale and regime size

Let us consider a time series of a dyadic length X_t, t= 1, …, N where N=512. Let $\{W_n : n = 0,\ldots, N-1\}$ represent the DWT coefficients as defined in Equation (2.4) in Chapter 2, W= $\mathcal{W}\boldsymbol{X}$, where W is a column vector of length $N = 2^J$ whose nth element is the nth DWT coefficient W_n and $\mathcal{W}$ is an $N \times N$ satisfying $\mathcal{W}^T\mathcal{W} = I_N$. Note that the DWT coefficients are proportional to differences in adjacent averages of X_t on a scale by scale basis. As mentioned in Chapter 2, the number of coefficients at each level is exactly half that of the next higher level, so there are 256 coefficients at level 1 (λ_1), 128 coefficients at level 2 (λ_2), and so on. The elements of $\mathcal{W}$ can be arranged such that the first $N/2$ DWT coefficients are associated with unit scale changes, the next $N/4$ coefficients, with changes on a scale of two, and so forth until we come to the coefficients W_{N-4} and W_{N-3}, which are associated with changes on a scale of $N/4$, the coefficients W_{N-2} are associated with a change on a scale $N/2$, and finally W_{N-1} is proportional to the average of all the data. There are thus exactly $N/(2\lambda_j)$ coefficients in the DWT associated with changes on scale λ_j, where $\lambda_j = 2^{j-1}$ for $j = 1,\ldots, J$. So we can write

$$W = \begin{bmatrix} W_0 \\ \vdots \\ W_{255} \\ \vdots \\ W_{383} \\ \vdots \\ W_{511} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2}\left(X_1 - X_0\right) \\ \vdots \\ 1/\sqrt{2}\left(X_{511} - X_{510}\right) \\ \vdots \\ 1/2\left(X_{511} + X_{510} - X_{509} - X_{508}\right) \\ \vdots \\ 1/16\sqrt{2}\left(X_{511} + X_{510} + \cdots + X_{255} + X_{254} \cdots + X_0\right) \end{bmatrix} \tag{6.10}$$

Let

$$\overline{X}_t(\lambda) = \frac{1}{\lambda}\sum_{l=0}^{\lambda-1} X_{t-l} \tag{6.11}$$

be the average of the $\lambda \in \mathbb{Z}^+$, contiguous data values with indices from $t-\lambda+1$ to t. Note that $\overline{X}_t(1) = \overline{X}_t$, which can be regarded as a single point average, and $\overline{X}_{N-1}(N) = \overline{X}$, which is the sample average of all N data points.

Using the definition (6.10), we can write the W_n's as

$$\begin{Bmatrix} W_0 = \frac{1}{\sqrt{2}}\left[\overline{X}_1(1) - \overline{X}_0(1)\right], \ldots, W_{255} = \frac{1}{\sqrt{2}}\left[\overline{X}_{511}(1) - \overline{X}_{510}(1)\right] \\ W_{256} = \left[\overline{X}_3(2) - \overline{X}_1(2)\right], \ldots, W_{383} = \left[\overline{X}_{511}(2) - \overline{X}_{509}(2)\right] \\ \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ W_{510} = 16\sqrt{2}\left[\overline{X}_{511}(256) - \overline{X}_{255}(256)\right] \\ W_{511} = 32\sqrt{2}\left[\overline{X}_{511}(512)\right] \end{Bmatrix} \tag{6.12}$$

Additionally, there is one coefficient W_{N-1} associated with an average at scale N. The rows of $\mathcal{W}$ that produce the wavelet coefficients for a particular scale are circularly shifted versions of each other, which are seen in Equation (6.12). Each wavelet coefficient at each scale is also localised in time. In this illustration, W_0 involves times points t = 0 and 1, whereas W_{256} involves times points t = 0, 1, 2 and 3.

Let us now decompose the elements of the vector W into J+1 subvectors as defined in Equation (2.12). The first J vectors are denoted by $\boldsymbol{W}_j, j = 1,...,J,$ where the jth subvector contains all of the DWT coefficients for scale λ_j. Note that $\boldsymbol{W}_j$ is a column vector with $N / 2^j$ elements.

$$W = \mathcal{W}X = \begin{bmatrix} \mathcal{W}_1 \\ \mathcal{W}_2 \\ \vdots \\ \mathcal{W}_J \\ \mathcal{V}_J \end{bmatrix} X = \begin{bmatrix} \mathcal{W}_1 X \\ \mathcal{W}_2 X \\ \vdots \\ \mathcal{W}_J X \\ \mathcal{V}_J X \end{bmatrix} = \begin{bmatrix} \boldsymbol{W}_1 \\ \boldsymbol{W}_2 \\ \vdots \\ \boldsymbol{W}_J \\ \boldsymbol{V}_J \end{bmatrix} \tag{6.13}$$

The final subvector is denoted as V_J and contains just the scaling coefficient W_{N-1}. We can then write $W = [\boldsymbol{W}_1, \boldsymbol{W}_2, ..., \boldsymbol{W}_J, \boldsymbol{V}_J]^T$. When $N = 2^J = 512$ so that J=9, we have by changing the regime size, we will implement a DWT analysis with N=512 to investigate which wavelet scale (level) detects the change point, the latter is termed the optimal scale in Section 6.3.

$$\left\{ \begin{array}{l} \boldsymbol{W}_1^T = [W_0, W_1, ..., W_{255}] \Rightarrow Scale\ 1 (\lambda_1 : d_1 = 256 \text{ coefficients}) \\ \boldsymbol{W}_2^T = [W_{256}, W_{257}, ..., W_{383}] \Rightarrow Scale\ 2\ (\lambda_2 : d_2 = 128 \text{ coefficients}) \\ \boldsymbol{W}_3^T = [W_{384}, W_{385}, ..., W_{447}] \Rightarrow Scale\ 4\ (\lambda_3 : d_3 = 64 \text{ coefficients}) \\ \boldsymbol{W}_4^T = [W_{448}, W_{449}, ..., W_{479}] \Rightarrow Scale\ 8\ (\lambda_4 : d_4 = 32 \text{ coefficients}) \\ \boldsymbol{W}_5^T = [W_{480}, W_{449}, ..., W_{495}] \Rightarrow Scale\ 16 (\lambda_5 : d_5 = 16 \text{ coefficients}) \\ \boldsymbol{W}_6^T = [W_{496}, W_{497}, ..., W_{503}] \Rightarrow Scale\ 32 (\lambda_6 : d_6 = 8 \text{ coefficients}) \\ \boldsymbol{W}_7^T = [W_{504}, W_{505}, ..., W_{507}] \Rightarrow Scale\ 64 (\lambda_7 : d_7 = 4 \text{ coefficients}) \\ \boldsymbol{W}_8^T = [W_{508}, W_{509}, W_{510}] \Rightarrow Scale\ 128 (\lambda_8 : d_9 = 2 \text{ coefficients}) \\ \boldsymbol{W}_9^T = [W_{511}] \Rightarrow Scale\ 256 (\lambda_9 : d_9 = 1 \text{ coefficient}) \\ \boldsymbol{V}_9^T = [W_{511}] \Rightarrow (s_9 = 1 \text{ coefficient}) \end{array} \right\} \tag{6.14}$$

6.2.3 The Shifting DWT (SDWT) and the MODWT

The motivation for using the MODWT is essentially to define a transform that will act as much as possible like the DWT, but does not suffer from the DWT's sensitivity to the choice of a starting point. This sensitivity is due to downsampling the outputs from the wavelet and scaling filters at each stage of the pyramid algorithm. Furthermore we modify the MODWT wavelet coefficients to link to our DWT shifting index, which will be used in the next section to detect a change point.

By way of a simple example, we consider N = 16 and the Haar wavelet, as follows:

$$\widetilde{W} = \widetilde{\mathcal{W}}X$$
$$= \left[\widetilde{\mathcal{W}}_1, \widetilde{\mathcal{W}}_2, \widetilde{\mathcal{W}}_3, \widetilde{\mathcal{W}}_4, \widetilde{\mathcal{V}}_4\right]^T \times \left[X_1, X_2, X_3, \ldots, X_{16}\right]^T$$

Next we have

$$\widetilde{\mathcal{W}}_1 = \tilde{h}_1 * \begin{bmatrix} \tilde{h}_1 & 0 & \ldots & 0 & 0 & -\tilde{h}_1 \\ -\tilde{h}_1 & \tilde{h}_1 & 0 & \ddots & 0 & 0 \\ 0 & -\tilde{h}_1 & \tilde{h}_1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & \vdots \\ 0 & \ddots & 0 & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & -\tilde{h}_1 & \tilde{h}_1 \end{bmatrix} = \left(\tilde{h}_1\right)^2 \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & -1 \\ -1 & 1 & 0 & \ddots & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ldots & 0 & -1 & 1 \end{bmatrix}$$

where $\tilde{h}_1 = \frac{1}{\sqrt{2}}$, and the MODWT wavelet coefficients can then be written as

$$\widetilde{W}_{1,t} = \widetilde{\mathcal{W}}_1 X = \left(\tilde{h}_1\right)^2 \left[X_1 - X_{16}, X_2 - X_1, X_3 - X_2, X_4 - X_3, X_5 - X_4, \ldots, X_{15} - X_{14}, X_{16} - X_{15}\right]^T$$

So $\widetilde{W}_{1,1} = \left(\tilde{h}_1\right)^2 \left[X_1 - X_{16}\right]$, and $\widetilde{W}_{1,16} = \left(\tilde{h}_1\right)^2 \left[X_1 - X_{16}\right]$ and so on.

Next we use the DWT and then shift the DWT wavelet coefficients by a unit scale using $W = \mathcal{W}X$. It then follows that

$$W_{1,t} = \mathcal{W}_1 X = h_1 \left[X_2 - X_1, X_4 - X_3, X_6 - X_5, \ldots, X_{14} - X_{13}, X_{16} - X_{15}\right]^T$$

where $\mathcal{W}_1$ is

$$\mathcal{W}_1 = \begin{bmatrix} -h_1 & h_1 & 0 & 0 & \ldots & 0 \\ 0 & -h_1 & h_1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \vdots & \vdots & 0 & 0 & \cdots & 0 \\ 0 & \ldots & \ldots & 0 & -h_1 & h_1 \end{bmatrix} \text{ and } h_1 = \frac{1}{\sqrt{2}}. \qquad (6.16)$$

After one backward shifting of the series X, the DWT wavelet coefficients are $W_{1,t}^{Shift1} = \mathcal{W}_1 X_{shift1} = h_1 \left[X_3 - X_2, X_5 - X_4, X_7 - X_6, ..., X_{15} - X_{14}, X_1 - X_{16} \right]^T$,

$$W_{1,t}^{shift0} = \frac{1}{\sqrt{2}} \begin{bmatrix} X_2 - X_1 \\ X_4 - X_3 \\ .. \quad ... \quad ... \\ X_{14} - X_{13} \\ X_{16} - X_{15} \end{bmatrix}, \; W_{1,t}^{Shift1} = \frac{1}{\sqrt{2}} \begin{bmatrix} X_3 - X_2 \\ X_5 - X_4 \\ ... \quad ... \quad ... \\ X_{15} - X_{14} \\ X_1 - X_{16} \end{bmatrix}. \qquad (6.17)$$

When we combine the row of $W_{1,t}^{shift0}$ (unshifted as in Equation (6.16)) and $W_{1,t}^{Shift1}$ (shifted as in Equation (6.17)), we then obtain the shifting DWT (SDWT) of λ_1, which can be expressed as

$$\text{SDWT= Combine } \left(W_{1,t}^{shift0}, W_{1,t}^{shift1} \right) = 1/\sqrt{2} \begin{bmatrix} X_2 - X_1 \\ X_3 - X_2 \\ \cdots \quad \cdots \\ \cdots \quad \cdots \\ \cdots \quad \cdots \\ X_{16} - X_{15} \\ X_1 - X_{16} \end{bmatrix} = \begin{bmatrix} X_2 - X_1 : sh0 d_{11} \\ X_3 - X_2 : sh1 d_{11} \\ X_4 - X_3 : sh0 d_{12} \\ X_5 - X_4 : sh1 d_{12} \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ \cdots \quad \cdots \quad \cdots \quad \cdots \\ X_{14} - X_{13} : sh0 d17 \\ X_{15} - X_{14} : sh1 d17 \\ X_{16} - X_{15} : sh0 d18 \\ X_1 - X_{16} : sh1 d18 \end{bmatrix},$$

where *sh0* is a non shift, and *sh1* is a unit shift, and so on.

SDWT then gives MODWT wavelet coefficients on a unit scale,

$$\text{SDWT}(W_{2,t})=\text{Combine}\left(W_{2,t}^{shift0},W_{2,t}^{shift1},W_{2,t}^{shift2},W_{2,t}^{shift3}\right)$$

$$=\left(\begin{bmatrix}X_4+X_3-X_2-X_1\\X_8+X_7-X_6-X_5\\X_{12}+X_{11}-X_{10}-X_9\\X_{16}+X_{15}-X_{14}-X_{13}\end{bmatrix},\begin{bmatrix}X_5+X_4-X_3-X_2\\X_9+X_8-X_7-X_6\\X_{13}+X_{12}-X_{11}-X_{10}\\X_1+X_{16}-X_{15}-X_{14}\end{bmatrix},\begin{bmatrix}X_6+X_5-X_4-X_3\\X_{10}+X_9-X_8-X_7\\X_{14}+X_{13}-X_{12}-X_{11}\\X_2+X_1-X_{16}-X_{15}\end{bmatrix},\begin{bmatrix}X_7+X_6-X_5-X_4\\X_{11}+X_{10}-X_9-X_8\\X_{15}+X_{14}-X_{13}-X_{12}\\X_3+X_2-X_1-X_{16}\end{bmatrix}\right)$$

$$=\begin{bmatrix}X_4+X_3-X_2-X_1:sh0d_{21}\\X_5+X_4-X_3-X_2:sh1d_{21}\\X_6+X_5-X_4-X_3:sh2d_{21}\\X_7+X_6-X_5-X_4:sh3d_{21}\\X_8+X_7-X_6-X_5:sh0d_{22}\\X_9+X_8-X_7-X_6:sh1d_{22}\\X_{10}+X_9-X_8-X_7:sh2d_{22}\\X_{11}+X_{10}-X_9-X_8:sh3d_{22}\\\cdots\quad\cdots\quad\cdots\quad\cdots\\\cdots\quad\cdots\quad\cdots\quad\cdots\\\cdots\quad\cdots\quad\cdots\quad\cdots\\\cdots\quad\cdots\quad\cdots\quad\cdots\\X_{16}+X_{15}-X_{14}-X_{13}:sh0d24\\X_1+X_{16}-X_{15}-X_{14}:sh1d24\\X_2+X_1-X_{16}-X_{15}:sh2d24\\X_3+X_2-X_1-X_{16}:sh3d24\end{bmatrix},\quad\text{and the MODWT is }\widetilde{W}_{2,t}=\begin{bmatrix}X_1+X_{16}-X_{15}-X_{14}\\X_2+X_1-X_{16}-X_{15}\\X_3+X_2-X_1-X_{16}\\\cdots\quad\cdots\quad\cdots\quad\cdots\\\cdots\quad\cdots\quad\cdots\quad\cdots\\X_{13}+X_{12}-X_{11}-X_{10}\\X_{14}+X_{13}-X_{12}-X_{11}\\X_{15}+X_{14}-X_{13}-X_{12}\\X_{16}+X_{15}-X_{14}-X_{13}\end{bmatrix}\tag{6.19}$$

$$\widetilde{W}_{1,t}=1/2\begin{bmatrix} X_1 - X_{16} \\ X_2 - X_1 \\ \cdots \quad \cdots \\ \cdots \quad \cdots \\ X_{15} - X_{14} \\ X_{16} - X_{15} \end{bmatrix}, \; t = 1, \ldots N. \tag{6.18}$$

For λ_1, the first value $\widetilde{W}_{1,1}$ $(=X_1 - X_{16})$ sends into the last row of Equation (6.18). Combining $\left(W_{1,t}^{shift0}, W_{1,t}^{shift1}\right)$ is useful to detect sudden change and to generate the location of sudden change using the shift index. The shifting is done on a scale with the shift length $= 2^j - 1$, $j = 1,2, \ldots, J$ in DWT. When we do this procedure for the next scale λ_2, we then have Equation (6. 19). For λ_2, if we send $\widetilde{W}_{2,1}$, $\widetilde{W}_{2,2}$, and $\widetilde{W}_{2,3}$ into $\widetilde{W}_{2,14}$, $\widetilde{W}_{2,15}$, and $\widetilde{W}_{2,16}$ in MODWT, it is the same as the SDWT. Thus the sequences of SDWT $W_{j,t}$ are obtained by MODWT.

The SDWT method is similar to MODWT which shows the strong correlation in the wavelet coefficients. It makes MODWT unsuitable for its use in the change point detection algorithms, which require the dataset (input values) to be independent. The DWT has poor resolution at higher scales and its lack of shift invariance. Hence the SDWT method turned our attention. The SDWT retains almost all the good properties of the MODWT, and what is more important, provides the time shift invariant the time series. By using this property of MODWT, we develop the SDWT which gives an accurate location of the change point in the time series.

6.3 Change point approach by the SDWT

6.3.1 Change point approach by wavelet thresholding

Donoho and Johnstone (1995b) propose the spatially adaptive properties of wavelet shrinkage estimators for change point detection. Smooth functions are characterized parsimoniously by defining the relation of parameters. Changing a fixed number bandwidth, the localised data points are able to characterize into wavelet coefficients by showing the sudden changes or discontinuities. From these motivations, we apply the Donoho and Johnstone's minimax wavelet threshold and also kernel smooth regression to detect a change point in a time series or function.
To investigate the performance of our proposed methods we consider, the square wave function with σ (=2) and added noise (SNR=1), given in Figure 6.1. Figure 6.2 results from applying the minimax thresholding and soft thresholding rule to a set of X_t, t = 1, ...512, of equally spaced

values from the square wave with σ=1 (top two plots) , σ=2 (bottom two plots), SNR=1 and various threshold values. For example, the minimax threshold η=2.048 this is suggested by Donoho and Johnstone (1994) was chosen for the series size 512. Figure 6.3 shows the kernel regression with square wave with σ=1 (top two plots), σ=2 (bottom two plots), SNR=1 and various bandwidth. The qualitative effect of bandwidth choice on kernel estimators is an important factor. A very small choice of bandwidth will localise the estimator too much and then the resulting estimator will consist of a lot of sharp bumps centred at the data values. Conversely, using a very large value of bandwidth will not allow sufficient localisation to occur, the resultant estimator will then smooth over any local features, completely obscuring the true picture. The bandwidth= 25 is chosen in example so as to minimize the mean squared error (≈ 0) between X_t and the kernel estimator.

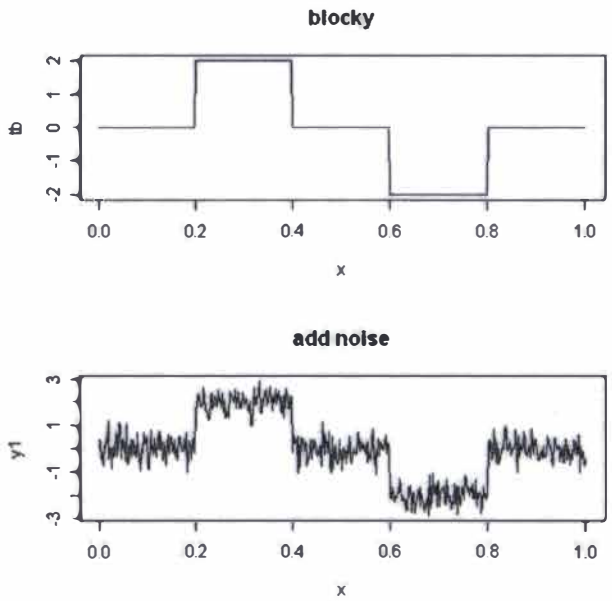


Figure 6.1 The square wave function with added noise.

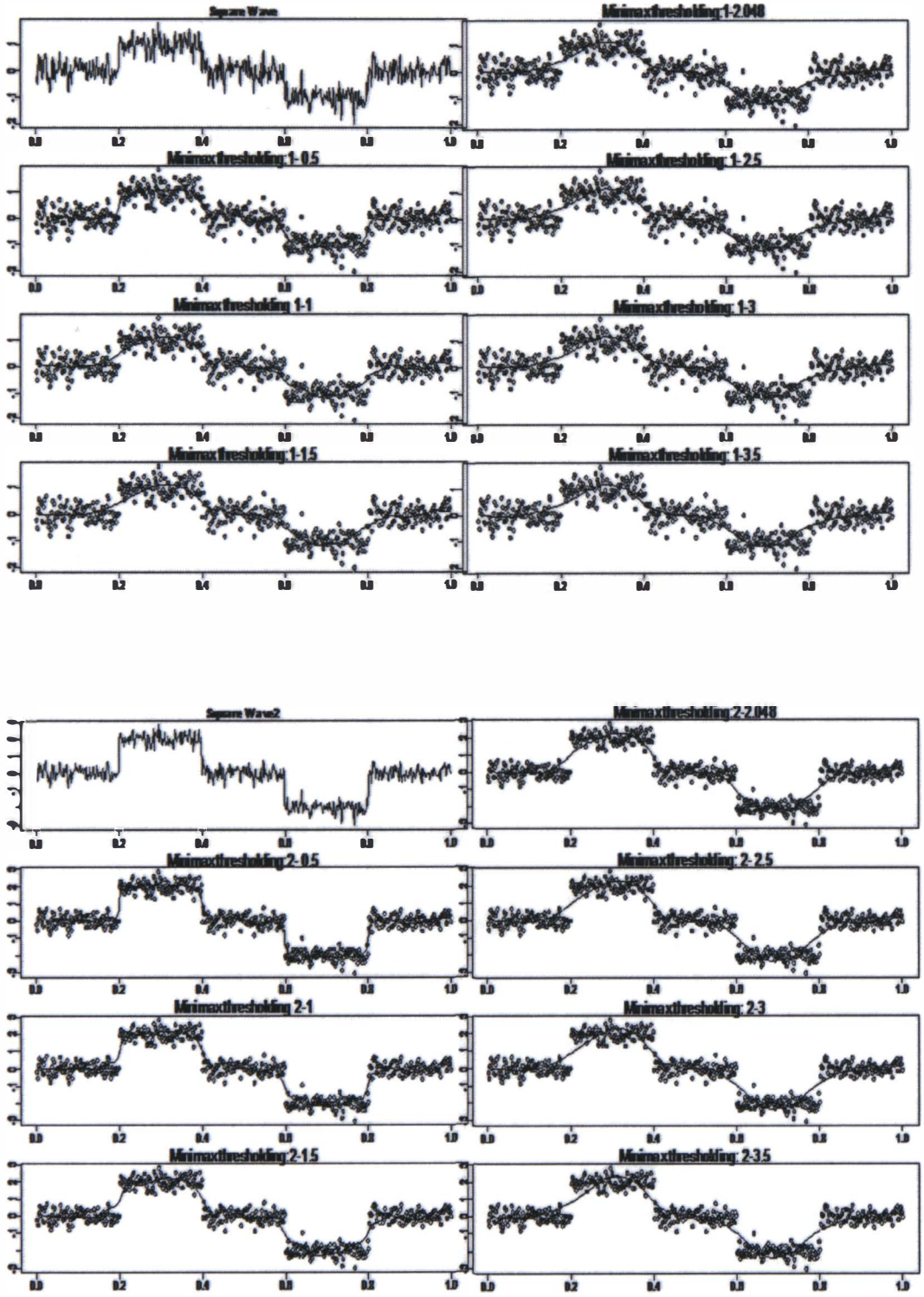


Figure 6.2 The square wave function and its minimax threshold estimator with various threshold values.

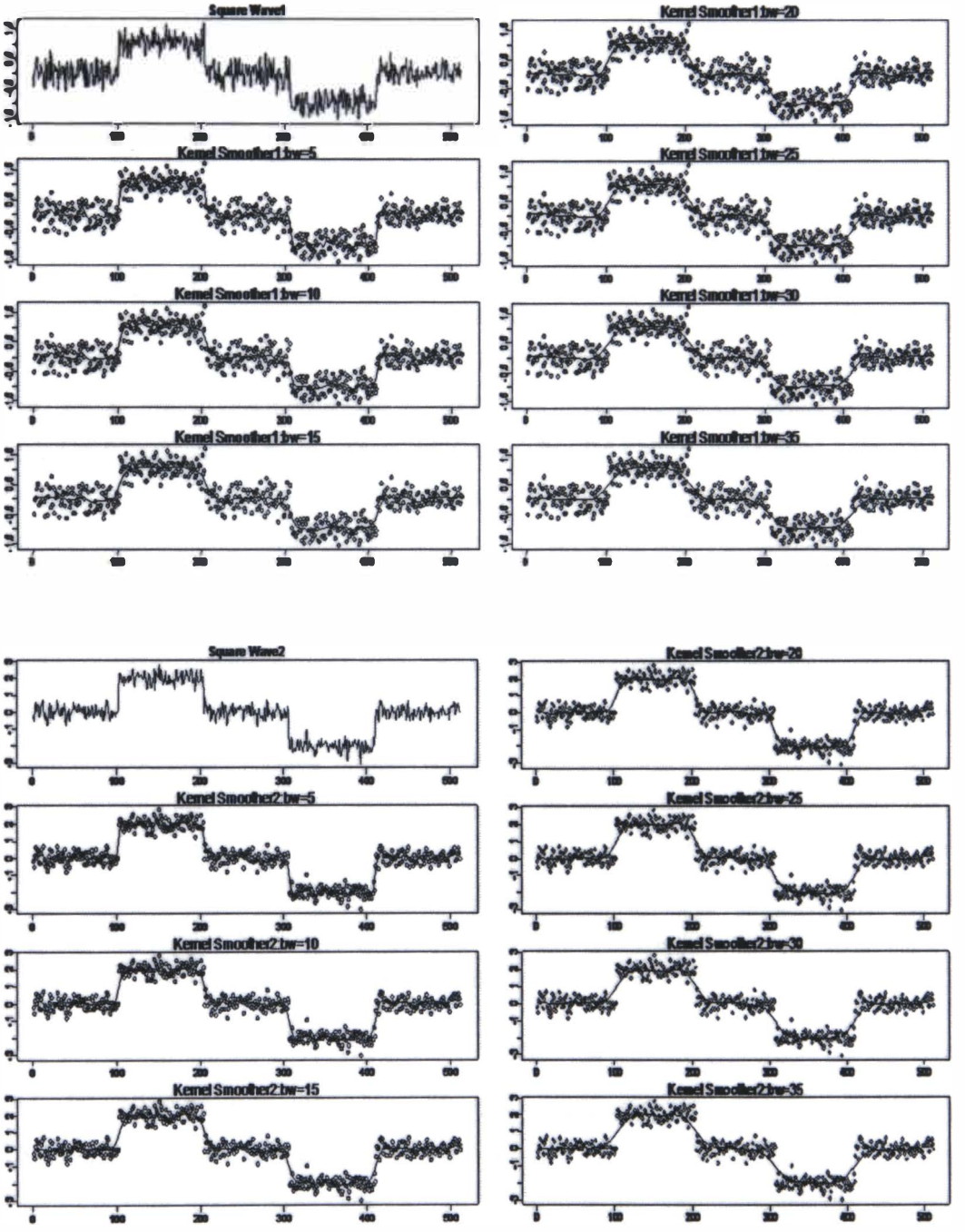


Figure 6.3 The square wave function and its kernel regression function with various bandwidth.

From Figure 6.2 and 6.3, we observe that minimax thresholding and the kernel regression method are not good enough to detect the change points in a square wave because of the smoothing information about the change point is split between the smooth and residuals, so provided less power for testing. Furthermore, there is a tendency to over-smooth a function around change points in practice, which has already been verified in several contexts (e.g., see of Hardle et al., 1998, Chapter 10).

The idea of using a Kolmogorov-Smirnov test statistic for wavelet coefficients was proposed previously by Ogden and Parzen (1996a) who were interested in applying change point methods to the problem of wavelet thresholding. Their work led to a threshold that is determined by the data on a scale-by scale basis. Additional references with respect to data dependent threshold selection include Abramovich and Benjamini (1995) and Ogden and Parzen (1996b).

Many smoothing technique are used to estimate of the nonparametric of regression models which were assumed as the regression function is smooth. Nevertheless, the smooth curve can contain a change point at which a jump or a sharp cusp occurs. Particularly the long time series or non-stationary time series, there are more possibilities to have change points. e.g.; Jaruskova (1997) propose that there is a chance of change points in the process of meteorological observations. Horváth and Kokoszka (1997) revise a class of change-point estimators specifically for the latter mean shift Gaussian model. Wang (2008) proposes several change-point estimators based on kernel and nearest neighbor methods for the random design model. On the other hand, the kernel regression estimates used by Wang (2005, 2008) are not good models for detecting the sharp cusps of a curve due to a single-bandwidth kernel estimate which tends to ignore some fine local details of the curve.

To deal with the change point problem we adopt the SDWT analytic approach. Next we demonstrate how the SDWT may be used to construct a statistical test for homogeneity of variance of the wavelet coefficients in an observed time series. The SDWT is used to detect changes like jumps and sharp cusps of signals embedded in Gaussian white noise.

6.3.2 Estimation of the wavelet variance

Allan (1966) gives a simple example of a wavelet variance estimator that measures the frequency stability for clocks and oscillators (Allan, 1966), later which is used to measure the wavelet variance by Percival and Walden (2000). Percival and Walden (2000) also defined that the wavelet variance is a succinct alternative to the power spectrum based on the Fourier transform,

yielding a scale-based analysis that is often easier to interpret the frequency-based spectrum. For more review of the wavelet variance see Percival and Walden (2000) and Gencay et al (2001).

We defined the definition of the wavelet variance in Section 2.2.3. We now consider estimation of the variance of the wavelet coefficients using some robust estimators for outlier detection, based on our SDWT. We use the estimation error to both determine the value of the influence function and to identify outliers, which are separated out, thus decreasing the effects of those points possessing large estimation errors. As the wavelet variance decomposes the variance of certain stochastic processes into sequences of wavelet coefficients that are associated with different time scales, it provides a very useful tool in the analysis of time series that fluctuate across a wide range of time scales.

The next section introduces a scale-based contamination series (noisy series) and describes robust estimation of the wavelet variance that can guard against such contaminations.

6.3.2.1 Robust estimators and the M-estimator of the wavelet variance

The sample standard deviation is sensitive to outliers or extreme values. Robust scale estimators are less affected by an arbitrarily large or small value. We compute several robust measures of scale, including the interquartile (IQR) range, Gini's mean difference (*Gini*), and the median absolute deviation about the median (MAD). In addition, we obtain estimates of the normal standard deviation σ derived from each of these measures (IQR^S, *Gini* and MAD), as described below.

The interquartile range is simply the difference between the upper and lower quartiles. For a normal population, σ can be estimated as IQR/1.34898 by Silverman (1986).

Next Gini's mean difference is computed as below

$$Gini = \frac{1}{\binom{n}{2}} \sum_{i<j} |x_i - x_j|.$$

For a normal population, the expected value of Gini's mean difference is $Gini = 2\sigma / \sqrt{\pi}$. When the data are a normal distribution, a robust estimator of σ is $Gini\sqrt{\pi} / 2$. This estimator has high efficiency relative to the usual sample standard deviation, and it is also less sensitive to the presence of outliers, when the underlying distribution is normal.

The median absolute deviation from the median (MAD) (Hampel, 1974), is computed as $\text{MAD}=\text{med}_i\left(\left|x_i-\text{med}_j\left(x_j\right)\right|\right)$. For a normal population, 1.4826MAD is an estimator of σ. For normal distributions due to the low efficiency of MAD, and it may not always be good for use it. Rousseeuw and Croux (1992, 1993) proposed two statistics as alternatives to the MAD to overcome the low efficiency. The first is

$$S_n=1.1926\text{med}_i\left(\text{med}_j\left(\left|x_i-x_j\right|\right)\right)$$

where the outer median (taken over *i*) is the median of the *n* medians of $\left|x_i-x_j\right|, j=1,2,\ldots,n$.

The second statistic is

$$Q_n=2.219\left\{\left|x_i-x_j\right|; i<j\right\}\binom{[n/2]+1}{2}.$$

Furthermore Croux and Rousseeuw (1992a, 1992b) refer to a bias-corrected statistic.

M-estimators are general estimators which are obtained as to minimise a certain robust properties of the data. Huber (1964) propose generalizing maximum likelihood estimation to the minimization of

$$\sum_{i=1}^{n}\rho\left(x_i,\theta\right), \tag{6.21}$$

where ρ is a arbitrary some function with certain properties. The solutions

$$\hat{\theta}=\arg\min_{\theta}\left(\sum_{i=1}^{n}\rho\left(x_i,\theta\right)\right) \tag{6.22}$$

are called M-estimators (Huber, 1981; Huber & Ronchetti, 2009; Mondal & Percival, 2009) which is a generalisation of the maximum likelihood estimators. The function ρ, and its derivatives, ψ, can be chosen to provide the estimator with minimizing the Equation (6.21). M-estimators give a solution, θ, which minimize Equation (6.22). Often this minimization is carried out by differentiation with respect to θ and solving. When this differentiation is possible, the M-estimator is said to be of ψ-type (Huber, 1981; Huber & Ronchetti, 2009). Otherwise, the M-estimator is said to be of ρ-type. In this study, we used Huber's ψ function (see Equation

(6.21)). More details are given in in Huber (1981), Wilcox (2005), Huber & Ronchetti (2009), Maronna et al. (2006), and Mondal & Percival (2009).

6.3.2.2 Confidence intervals for the M-estimators

A natural way of computing confidence intervals for μ_m, an M-measure of location, is to estimate the standard error of $\hat{\mu}_m$ with $\hat{\sigma}_m$ as a form of $\left(\hat{\mu}_m \pm \alpha\sigma_m\right)$. Where α is determined so that the probability coverage is exactly 1-α. We have that the distribution of wavelet coefficients of each scale are normally distributed with mean zero and variance as follows, N (0, robust estimators or M-estimator) with $\hat{\mu}_m = \bar{X}_m$. Therefore, we have $\left(0 \pm 1.96\hat{\sigma}_m\right)$ with α=0.5.

6.3.3 Testing the homogeneity of the wavelet variance

Given that the length of the wavelet filter L is sufficient to eliminate any possible polynomial trends, we have

$$\boldsymbol{W}_{j,t},\ t=L_j',\ldots,N_j-1,\ \text{where } L_j' = \left[\left(L-2\right)\left(1-\frac{1}{2^j}\right)\right],$$

is a nonboundary sequence of scale λ_j SDWT coefficients, and $N_j' = N_j - L_j' > 0$, $L_j' > 0$. We assume from the decorrelating property of the SDWT, $E(\boldsymbol{W}_{j,t}) = 0$ and Var $(\boldsymbol{W}_{j,t}) = \sigma_t^2\left(\lambda_j\right)$. We now will use these wavelet coefficients to construct a test statistic for the null hypothesis

$$H_0 : \sigma^2_{L_j'}\left(\lambda_j\right) = \sigma^2_{L_j'+1}\left(\lambda_j\right) = \ldots = \sigma^2_{\hat{N}-1}\left(\lambda_j\right)$$

Robust etimators of the wavelet variance are used (described in Section 6.3.2.1) as a test statistic, named **RE**, say, that can discriminate between the null hypothesis and a variety of alternative hypotheses, such as

$$H_1 : \sigma^2_{L_j'}\left(\lambda_j\right) = \ldots = \sigma^2_{k_1}\left(\lambda_j\right) \neq \sigma^2_{k_1+1} \ldots \sigma^2_{k_t} \neq \sigma^2_{k_t+1} \ldots = \sigma^2_{\hat{N}-1}\left(\lambda_j\right),$$

where t is the unknown number of change points and $1 \le k_1 < k_2 < \ldots < k_t < k_{t+1} < \ldots < N-1$ are the unknown locations of the change points. The critical values of the statistic **RE** are derived from robust estimators and M-estimators, described above, that are used to estimate the wavelet variance used in testing the null hypothesis.

To test the null hypothesis above against the alternative with unknown k_0

$$H_1 : \sigma^2_{L'_j}(\lambda_j) = \sigma^2_{L'_j+1}(\lambda_j) = \ldots = \sigma^2_{k_0} \neq \sigma^2_{k_0+1} \ldots = \sigma^2_{\hat{N}-1}(\lambda_j).$$

We propose the following procedure by which to establish a significant change point at (k_0+1). First compute the variance sequence simultaneously and also compute those related to pairs of subsequences on each side of k_0 (i = 1, ... , k_0 and i = k_0 + 1, .. , N, 1 < k_0 < N-1). Secondly, compare these values. Conclude that a significant change point exists at position (k_0+1), if the variance of the subsequences is lower than the variance of the whole sequence.

6.3.3.1 Location of a Change Point (LCP)

Once an observed time series has been tested and the null hypothesis rejected at some scale λ_j, we have succeeded in detecting a significant change in wavelet variance somewhere in the series. It is noteworthy where each wavelet coefficient is associated with each t regardless of scale, the location of a wavelet variance change may be associated with a specific observation in the original time series.

Although in practice wavelet coefficients $\boldsymbol{W}_j$ are computed using a pyramid algorithm, it is of theoretical interest to note that we could obtain their elements directly from $\boldsymbol{X}$ via

$$W_{j,t} = \sum_{l=0}^{L_j-1} h_{j,l} X_{2^j(t+1)-l-1 \bmod N},$$

where $h_{j,l}$ is the jth level equivalent wavelet filter, each having width $L_j = (2^j - 1)(L-1)+1$. Then the location of the change point (LCP) is

$$\mathbf{LCP} = \left[2^j \times \mathrm{LMWC}\right] + \mathrm{S{\cdot}I} - (2^j - 1)$$

where LMWC denotes the Location of Maximum Wavelet Coefficient defined as $\mathrm{LMWC}=\max\left[\boldsymbol{W}_{j,t}\right], j = 1,\ldots,J$ for each λ_j, S.I (λ_j) is the index which gives the maximum |WC| through the SDWT from 1 to $\left(\frac{N}{2^j} - 1\right)$ on each scale, and the last term $(2^j - 1)$ is scale adjust index for the boundary condition. If the procedure with no shift with the first wavelet coefficient (WC) gives the largest |WC|, then one should just follow the ordinary calculation of the location of the WC.

Let X_t, t= 1, …, 512 be the square wave which has the first regime size 16 with break size 2 and varying SNR as given in Figure 6.4. From Figure 6.4, when we add more noise to the function, it is hard to find the change point and if the regime size is smaller, it is harder to find the change point.

Before implementing the algorithm for finding the LCP, an example of finding the optimal shift index is shown in Figure 6.4 using the square wave (the top plot in the Figure 6.3). When we compare the wavelet coefficients scale by scale, the wavelet coefficient in level 5 (d_5) attains the greatest value from the choice based on d_1 to d_7. Level 5 (d_5) is the optimal scale that shows the exact location of the change point for this square wave example, this commensurate with the length of the regime of 2^{5-1}=16 (i.e. half of the wavelet interval, where the two level Haar wavelet changes level. The optimal level is clearly that with an wavelet interval length double the length of the regime length.

Next we use simulations (100 simulations) to illustrate the detection properties of our proposed SDWT algorithm and also to indicate the optimal scale. Figure 6.5 shows that the location of the change point on a scale-by-scale basis, with N= 512, R_1=16 for the short regime as an example; with the value of σ chosen to give a signal-to-noise ratio from 0.5 to 3.

The X-axis in Figure 6.6 is a wavelet scale from level 1 to 9, Y-axis is the simulation number, and Z-axis represents the LCP. Through 100 simulations, we attempt to find the optimal scale that provides the location of change point. By varying SNR, we can easily find the optimal scale 5, which is true since the change point is at timepoint 17, therefore the regime length is 16 and so the wavelet is on scale 5. Therefore LCP influences the optimal scale.
Now by using the M-estimator technique for estimating the variance, we check which wavelet coefficients are significant on scale-by-scale basis using the **RE** statistic. Figure 6.7 presents the signature of wavelet coefficients from the series in Figure 6.6 using the M-estimator's upper tail critical value given in Table 6.1 (see Section 6.3.2.2). Figure 6.7 shows the most significant wavelet coefficients are in λ_5 (=$d5$) (a maximum number of points above the critical (horizontal) line). We thus conclude that scale 5 is the optimal scale by which to detect a CP. This is in agreement with the results from the LCP method that also found scale 5, to be the optimal scale by which to detect a change point (Figure 6.6) (at 17). From Figures 6.6 and 6.7, we conclude that our SDWT in choosing the optimal scale by which to detect a change point when we have multi scales. Indeed we do not need then to test all wavelet scales to detect the change point.

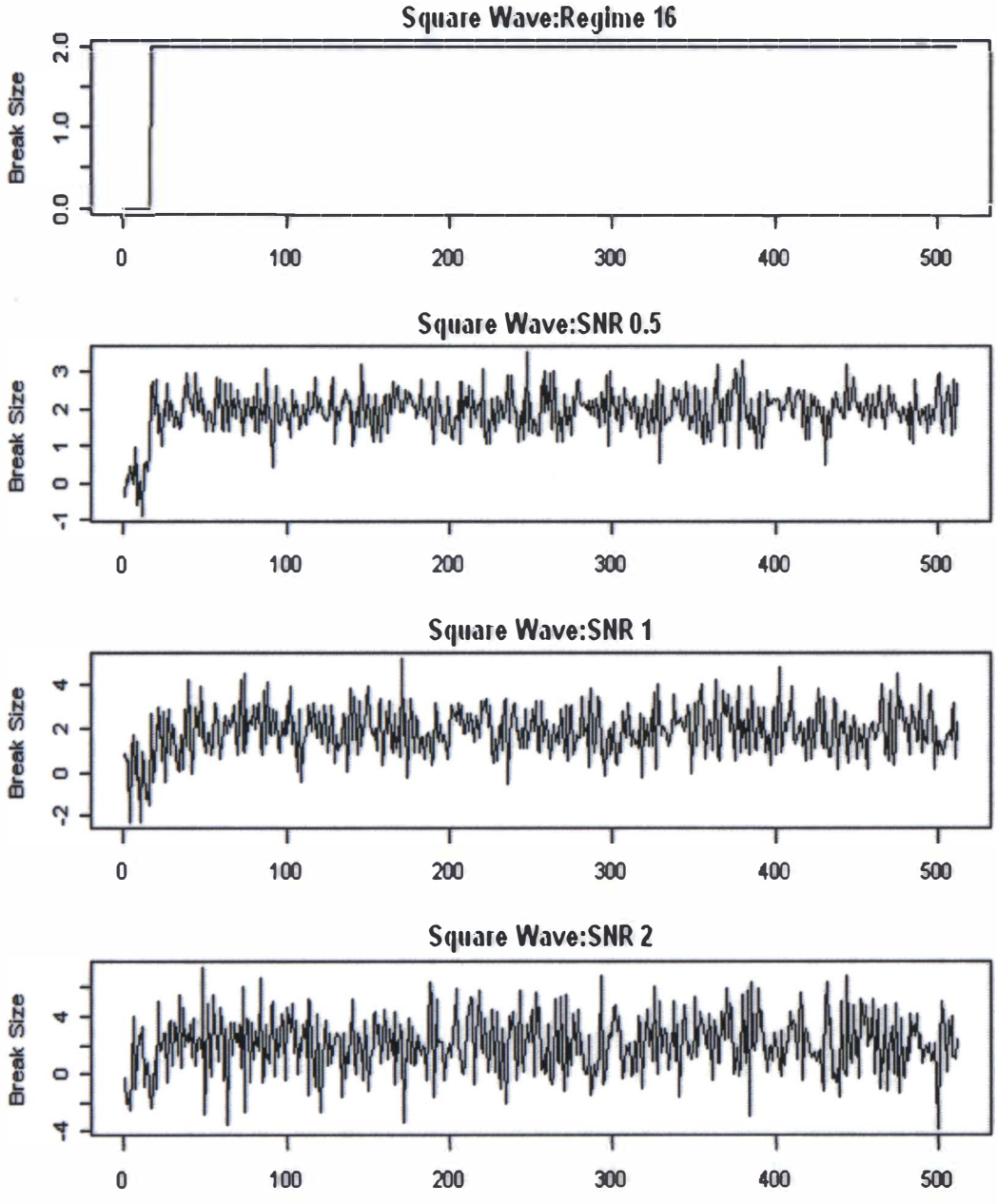


Figure 6.4 The quare wave for X_t with the first regime size (R_1): 16 and the second regime size (R_2): 496 and varying SNR

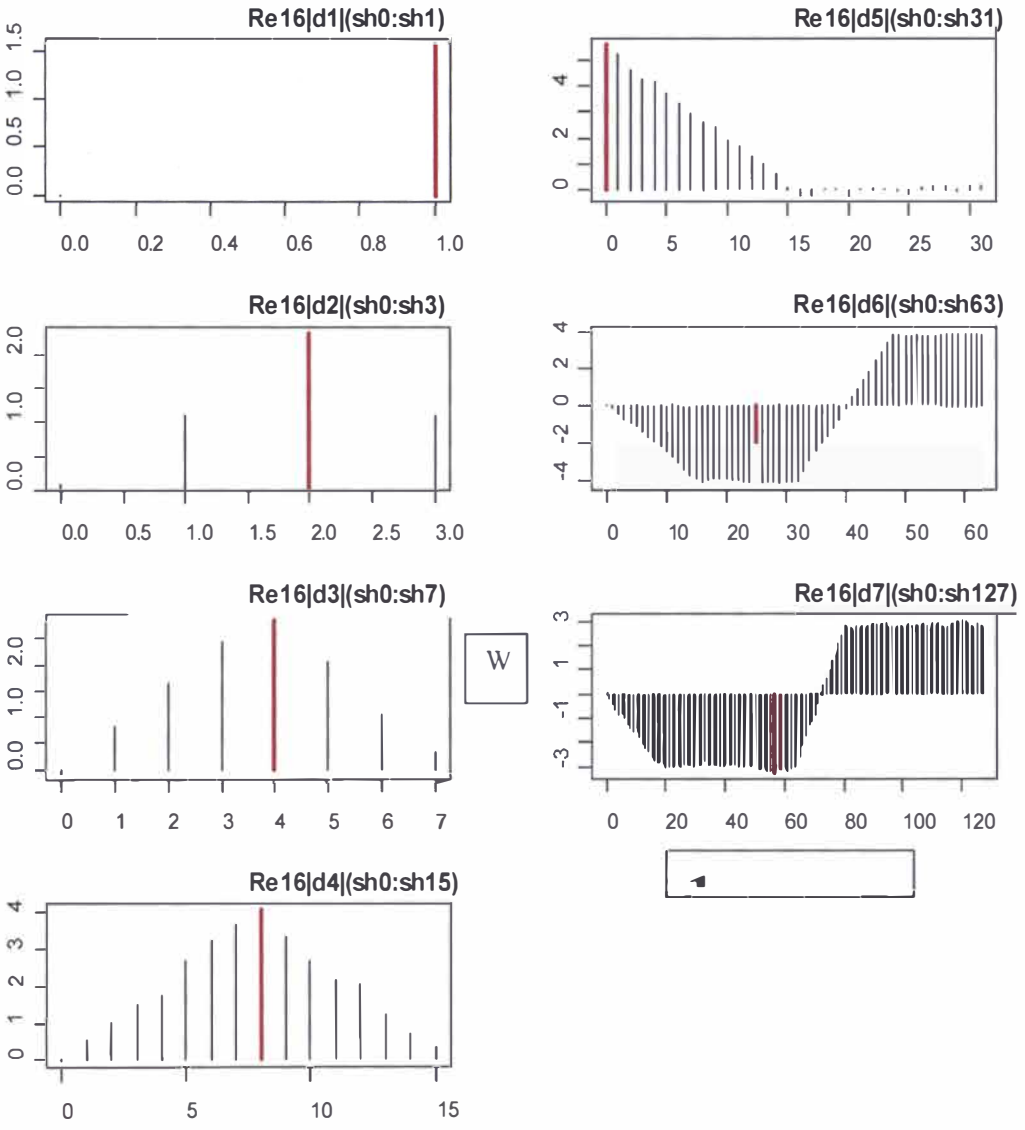


Figure 6.5 An example of the shifting index using DWT absolute coefficients for X_t with first regime size (R_1) 16 and $R_2 = 496$. The thick red line shows the largest absolute wavelet coefficient value on each scale.

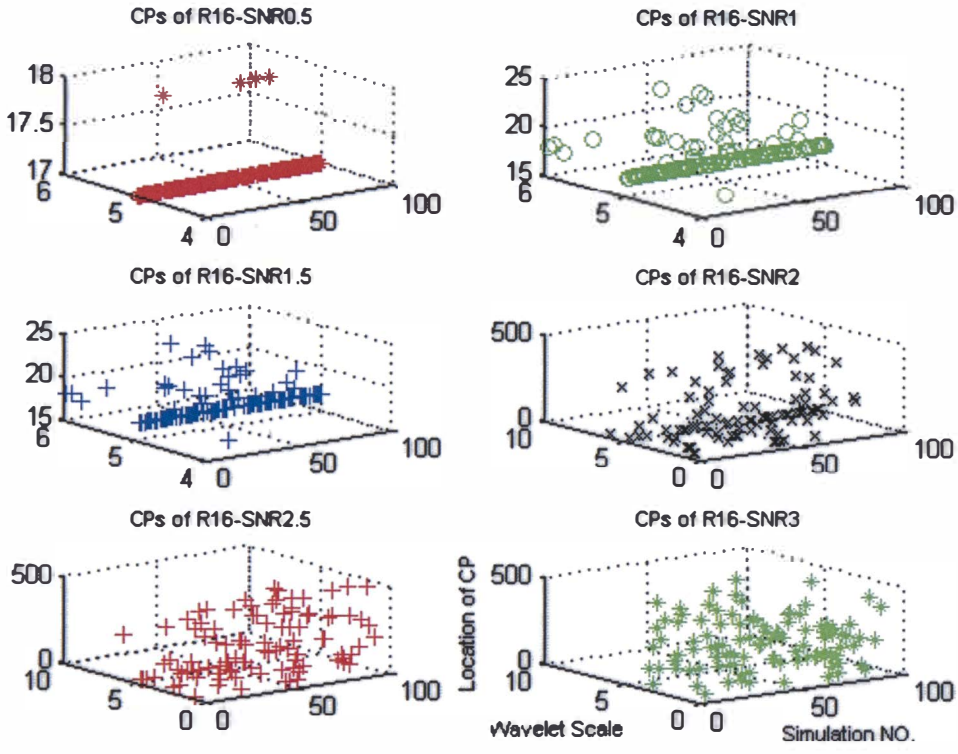


Figure 6.6 The location of the change point (CP) with N=512 series by using the LCP formula for the multi scales.

Table 6.1 The upper confidence interval value for the four estimators with α=0.5

	Estimator type - CI			
Scale	IQR-CI	Gini-CI	MAD-CI	M-CI
Scale 1	1.156	1.084	1.364	1.139
Scale 2	0.998	1.038	1.218	1.090
Scale 3	1.040	1.090	1.375	1.131
Scale 4	1.029	1.114	1.336	1.093
Scale 5	0.986	2.250	1.208	1.091
Scale 6	1.138	2.376	0.697	1.506

6.4 Simulations and Comparison

In this section we apply the methodology developed in the previous section called the SDWT and compare this change point detection method to two detection procedures: the atheoretical regression trees (ARTs) method (Cappelli, Penny, Rea, & Reale, 2007a; Cappelli, Rea, & Reale, 2007b), and the methodology proposed by Bai and Perron (1998, 2003), known as the BP method. The methods are compared via simulation.

In the following simulation study we take 9-regime sizes of 8, 12, 24, 48, 64, 96, 128, 192, and 256. The break size is fixed at 2. A set of 100 simulations were run with series of uncorrelated observations drawn from a standard Normal distribution with a single change point in the series (N=512). In essence the resultant series are square waves with an amplitude of break size and Gaussian noise of constant variance imposed on them as shown in Figure 6.4. These set of 100 simulations were thus run with a series of uncorrected observations drawn from a standard normal distribution with a single change point in the series N=512.

In spite of their sub-optimal results, ART has found broad application due to its computational efficiency. It allows ART to apply large data sets with relative simplicity. Probably the best known regression tree methodology is the Classification and Regression Tree (CART) of Breiman et al. (1984) to which the reader is referred for a detail description of ART (Cappelli, Penny, Rea, & Reale, 2007a; Cappelli, Rea, & Reale, 2007b).

Recently Bai and Perron (1998, 2003) proposed the procedure which based on Fisher's (1958) method of exact optimisation. While the BP produces an optimal partition of a time series, it takes long time to compute in application for the long data in practice like financial and geophysical time series. We show, however, in this section, that ART provides a practical alternative (to the SDWT) for locating structural breaks or change points in the means of the time series data. A poor predictive performance in ART is that the distribution of the response variable is not orthogonal or parallel to the predictor variables (See Figure 2.10 of Breiman et al., (1984), for an example).

Consider now the time series model to simulate y_t, $t = 1,...,n$

$$y_t = x_r + \varepsilon_t, \text{ where } r = 1,\cdots, R \text{ and } t = T_{r-1}+1,\cdots,T_r. \qquad (6.23)$$

where R is the number of regimes and (R-1) is the number of breakdates. Note we adopt the common notation that $T_o = 0$ and $T_R = T_n$. In this study we simulate two regimes with t =512 series and the value σ was chosen to give a signal to noise ratio ranging 0.5 to 3. Here the regime size also varies from 8 to 256 since all the level wavelet coefficients are subject to change.

First we apply wavelet methods using the R package `waveslim` (Whitcher et al., 2006) and `WaveThresh` (Nason, 2008) to the simulated data from model Equation (6.23). We now test for homogeneity of variance using the test statistic **RE** described in the previous Section 6.3.3. Essentially the procedure we follow is to test homogeneity of variance, using the DWT with a

Haar wavelet filter for the convenience and then shifting for the wavelet coefficients on a scale-by-scale basis. Shifting wavelet coefficients from 1 to $\left(\frac{N}{2^j}-1\right)$ and the values of **RE** where **RE** is defined to be in Section 6.3.3 for the scales, λ_j, are provided in Table 6.2 for the first regime size 64 case, along with critical values computed via RE (IQR, Gini, and MAD) and the M-estimator (M). Since RE for all scales is greater than any of the tabulated critical values, we can reject the null hypothesis of homogeneity of variance at all four method specific critical values. Since we reject the null hypothesis of homogeneity of variance for all scales of the time series, we can now estimate the location of the variance change point using the LCP (see Section 6.3.3.1). Results are given in Table 6.3.

Table 6.2 Results of testing the N=512 series for homogeneity of variance using the Haar wavelet filter and SDWT.

		Critical values with level 5%			
λ_j	RE statistic	M	IQR	Gini	MAD
1	3.655	0.249	0.998	1.958	2.162
2	2.622	0.367	1.052	1.964	2.205
3	4.275	0.406	0.933	1.901	2.115
4	6.167	0.927	0.942	2.059	2.187
5	7.958	0.751	0.993	2.113	2.115
6	9.921	2.457	0.841	1.936	2.108
7	11.898	5.599	2.787	4.718	2.337

Table 6.3 Location of change point (LCP) using wavelet method.

λ_j	LMWC	SI	LCP
1	32	1	65
2	16	2	65
3	8	4	65
4	4	8	65
5	2	16	65
6	1	32	65
7	1	0	65

Next we apply ART to the regression model in Equation (6.23). We used for the tree growing and pruning procedures is implemented using the `tree` (Ripley, 2005) contributed package in R.

From Figure 6.8 we have an ART driven LCP of 65.5 that is the root node, namely, the location of the change point. In 100 simulations we only take the first split (root node). Then we change the regime size and subsequently calculate the average of the location of the change point and their standard deviation.

Next we apply BP to the model (Equation (6.23)) with the minimum segment size set to 0.1 which is the fraction relative to the sample size (≈ 65/512). For the BP method we use the package `strucchange` (Zeileis et al., 2002) in R. The BP application gives a LCP of 65 for the first regime size of 64. From Table 6.3, Figure 6.8 and the application of the BP method we show that all three methods (SDWT, ART, and BP) find the same LCP of 65 for the given signal to noise ratio. In what follows we demonstrate however that the three methods may detect differing change points as we increase the signal-to-noise ratio of the series (see Figures 6.8, 6.9, 6.10, and 6.11).

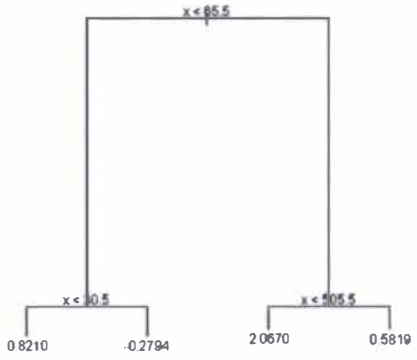


Figure 6.7 The location of change point found using ART (with N=512).

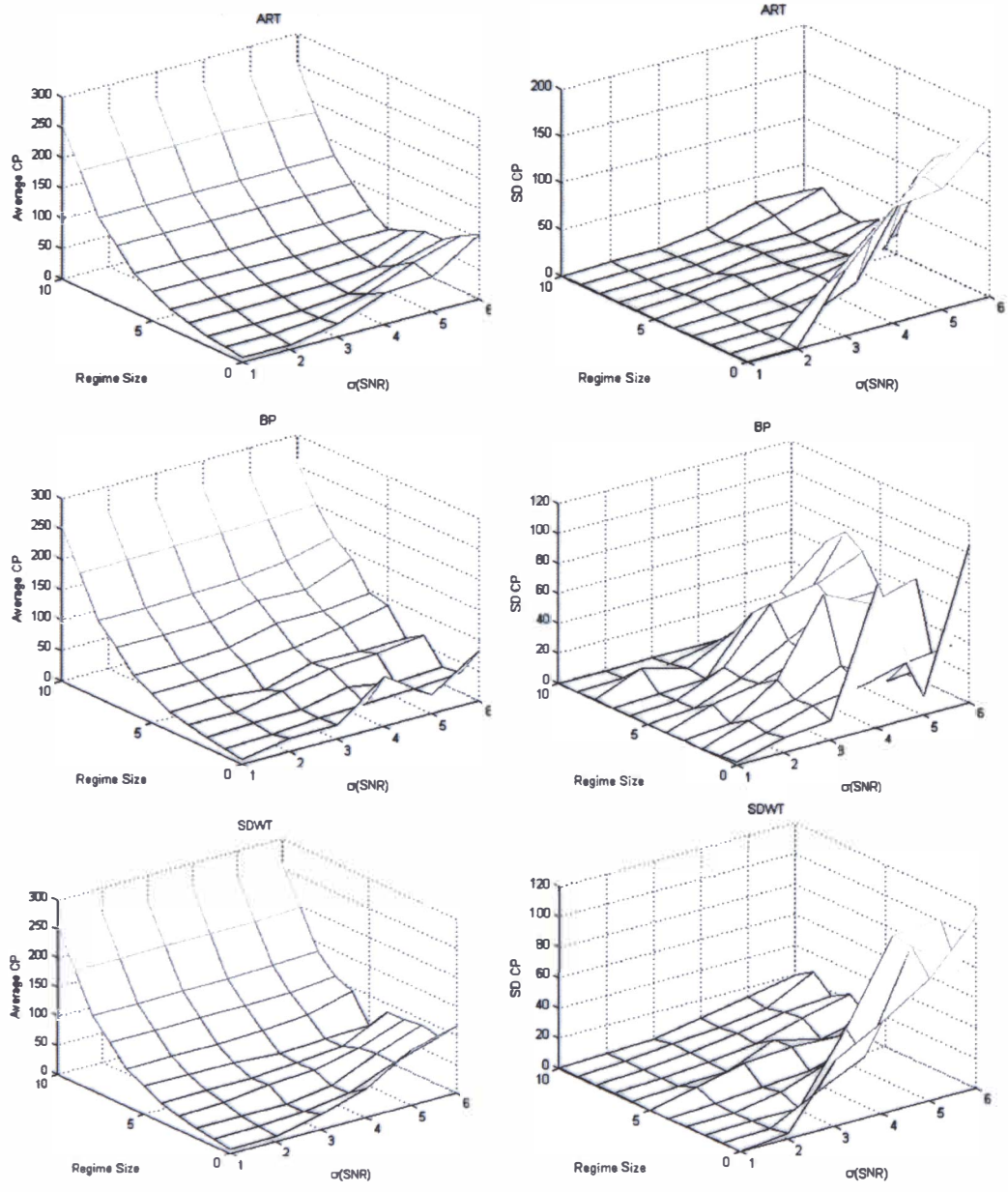


Figure 6.8 The average of the location of change point and standard deviation of the ART (Top), BP (middle), and SDWT (bottom) LCP detection methods: with varying regime size and σ of SNR.

Figure 6.9, presents the results for the ART, BP and SDWT analyses of a time series with a single break but with varying regime size and SNR. The results of three methods are very similar, as shown in Figure 6.9. As mentioned before we have 11 regimes and the regime size axis (of Figure 6.9) is non- linear in scale. When time series are short we note that the standard deviation of the location of change point via ART is very high but that this tendency gradually disappears when we reach a regime length of 48 (Figure 6.9). ART also had a tendency to over-fit for shorter regimes. This means that ART generates more than one change point.

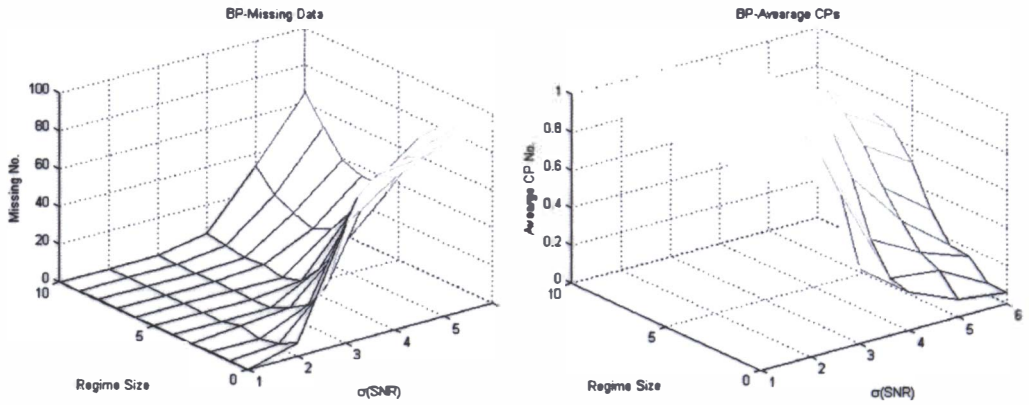


Figure 6.9 The missing number and an average number of the location of change point (LCP) for the BP method with varying regime size and SNR values.

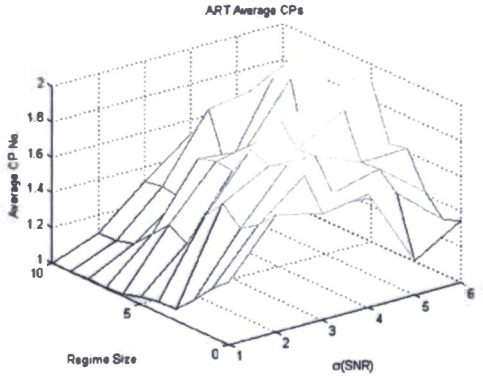


Figure 6.10 The average number of change points for the ART procedure, with varying regime size and SNR values.

The BP tended to under -fit for shorter regimes and high SNR (Figure 6.9). This means that BP cannot, in this example, detect any LCP (miss number of the location of change point; e.g. When BP generates only one value of LCP, then the missing value of LCP is 1 for two change points case)- a reason why the standard deviations of the change points are very high (Figure 6.9).

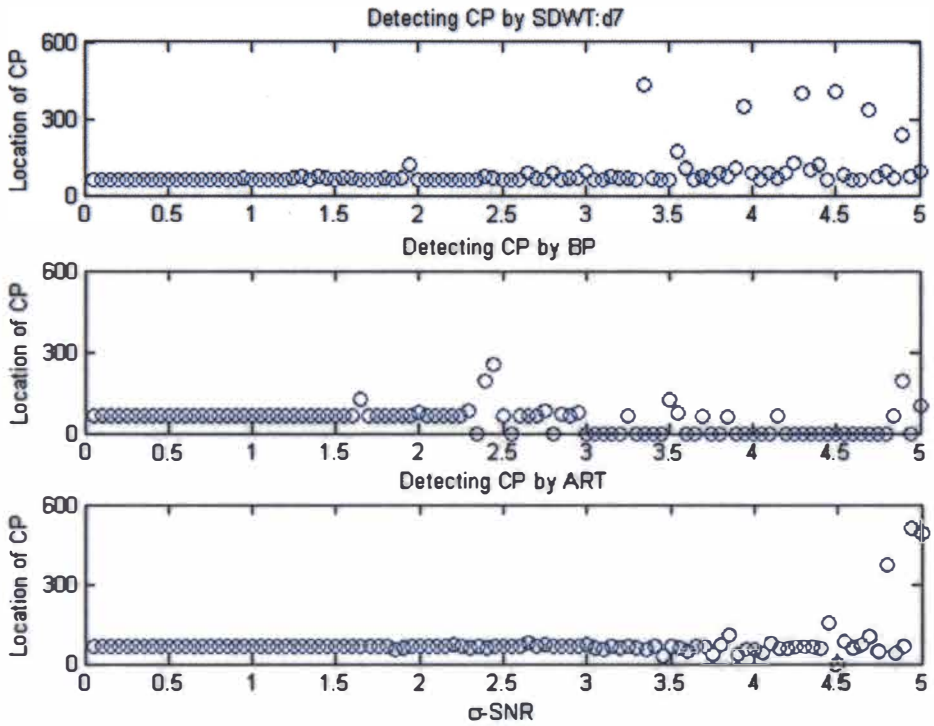


Figure 6.11 Comparison of the LCP across the optimal scale SDWT (SDWT d_7), ART, and BP methods with varying SNR for two regimes (R_1=65, first regime size, R_2= 446, second regime size) with N=512.

Figure 6.10 presents the results of detecting the location of change points (LCP) and the average number of LCP for the BP is presented in Figure 6.10. Figure 6.10 explains the high standard deviation of LCP in the Figure 6.9 since the missing data is increased when SNR is increased.

Figure 6.11 shows the average number of LCP for the ART with varying regime size and SNR. When the regime size is longer and the SNR is high, the number of the reported change point increases. This means that ART had a tendency to overfit for noisy longer regimes.

Figure 6.12 presents the comparison of the ART, BP and SDWT results for detecting a single change point at 65 on scale 7 with varying SNR (up to 5). The optimal scale for this series is λ_7 (=d_7) is. In BP method, the missing data were treated as 0 and we take the root node (first split) reported by the ART. By this simple comparison, ART is reliable even when the SNR is high and the BP method has a high variation with high SNR (see Figure 6.12).

6.5 Conclusions

The time–frequency localization of wavelets provides a natural motivation for using them in change point problems. The main goal here is the detecting the change points and finding their locations. Here we have developed a new DWT based methodology by using shifting of the data series as a data driven nonparametric procedure for detecting a change point.

The MODWT does not have orthogonality in order to gain features the DWT possess. Thus the wavelet and scaling coefficients must be rescaled in order to retain the variance preserving property of the DWT. MODWT seemed a good tool to analyze a time series because of its ability to localize, but it suffers from the presence of correlation in the transform. The correlation of the MODWT makes the transform not suitable for its use in the change point detection algorithms. The other hand, the MODWT is invariant to circularly shift the time series which property allows getting the location of the change point. The SDWT is actually combination of DWT (where shifting also retains the good properties of DWT) and MODWT.
The simulations have enabled us to overcome the problems of existing change point detection methods like ART and BP. ART can be easily implemented with packages containing routines to grow and prune least squares regression trees. Although ART and BP have advantages, when the series are short the ART is very prone to over –fitting but this tendency gradually disappears as the series becomes longer. BP tends to under-fit for shorter series and noisy series. Because of the amount of computation time required for BP, it is not a practical tool for analysing long time series.

The SDWT transforms a time series into coefficients with shift index that reflects the change at various wavelet scales and time. The SDWT identifies the optimal scale at which the inhomogeneity based on the wavelet variance occurs. The optimal wavelet scale provides detailed information about the series. The SDWT also estimates the time at which the variance change point occurs. SDWT also has advantages of simplicity since it can be easily implemented with packages which take a very short time to run. By comparison, the BP procedure takes a long time to run. ART and SDWT are, both computationally faster than BP, and both the ART and SDWT based procedures perform similarly in terms of computational speed and accuracy. SWDT also has an added advantage in that it can also be easily generalized to find multiple change points. This by way of a binary segmentation procedure, by iterating the process in the two subsequences that surround the first detected change point. If a change is detected subsequently, one can split again and iterate the process until no more change points are found in any of the subsequences.

Chapter 7

7 Conclusions

7.1 Overview of the Study

Overall, the research presented in this thesis shows that wavelets are appropriate tools in problems of signal processing, denoising, regression, and density estimation. The details are summarized in subsequent sections.

The wavelet methods advocated, developed and used in this thesis (Chapters 2, 3, 4 and 6) are primarily based on the discrete wavelet transform (DWT) and wavelet thresholding. The discrete wavelet transform (DWT) exhibits features that vary in both time and frequency. The DWT and the maximal overlap discrete wavelet transform (MODWT) are shown to be useful for the creation of informative descriptive statistics for both univariate and bivariate time series analysis.

The DWT, as a concept of time-frequency localization, is an effective statistical tool for re-expressing a time series in terms of coefficients that are associated with a particular time and a particular dyadic scale 2^{j-1}. These coefficients are fully equivalent to the original series in that one can perfectly reconstruct a time series from its DWT coefficients. The DWT allow us to partition the energy in a time series into subcomponents or details that are associated with different scales and times. This energy decomposition is thus very close to the statistical technique known as analysis of variance. Since a set of wavelet coefficients completely describes a time series, say, a graphical display of these DWT coefficients then potentially provides a good diagnostic plot for data analysis. Such a plot gives a good description of where significant change is taking place in the time series. The location of the abrupt jumps in the time series can thus be plotted as relatively large coefficients. This fact motivated the development of our DWT variant methods in Chapter 6, by which to detect a change point in time series data.

In this thesis, the DWT methods developed have been applied to modeling the agitation and sedation (A-S) time series profiles of patients in ICU whose A-S time series are of significant length (range [3,001-25,261] time points in minutes) (Chapters 2 and 3). In Chapter 3 a density estimation approach via wavelet smoothing was developed. This and novel wavelet density metrics were used to assess the validity of the deterministic dynamic A-S models (of the ICU data, Chapter 2) against empirical data. The results are compared to those obtained using the DWT and

MODWT based methods of Chapter 2. The approach in Chapter 3, based on wavelet smooth regression and its density estimation, are shown to yield excellent visual graphical assessment tools, in addition to numerical wavelet based metrics by which to statistically assess the compatibility between an ICU patient's simulated (deterministically modelled) and their true observed (recorded) A-S time series data.

Wavelet thresholding is used and compared with independent component analysis (ICA) as alternative denoising methods in Chapter 4, specifically applied to denoising a patient's SPECT brain image data, which was part of the neuroinformatics study of Turner et al. (2003) (Chapter 5). ICA methods are then implemented for denoising all the cerebral function data of Turner et al. (2003), presented as brain images (at a voxel by voxel basis). This was performed as a pre-processing step for the creation of statistical parametric maps by which to model personality with respect to brain function (Chapter 5) as non-linear models.

The major conclusions of the thesis are presented on a chapter by chapter basis in sections 7.2-7.4.

7.2 Wavelet methods to assess agitation and sedation in ICU patients

The primary goal of ICU sedation is to control a patient's agitation, while preventing over-sedation and over use of drugs. Agitation management via effective sedation management is a vital and fundamental activity in the ICU. However, in clinical practice a lack of understanding of the underlying dynamics of agitation and sedation, combined with a lack of subjective assessment tools, makes effective and consistent clinical agitation management problematic. Current clinical practice utilizes subjective agitation and sedation assessment scales, combined with medical staff's experience and intuition, to deliver appropriate sedation, on a patient by patient basis. However, such an approach typically leads to largely continuous infusions, which lack a bolus-focussed approach, and often result in either over sedation, or insufficient sedation. Recent studies have also highlighted the cost and healthcare benefits of drug delivery protocols based on validated agitation-sedation assessment scales. The work in Chapter 2 and Chapter 3 develops novel quantitative models and measures based on wavelets for the analysis of agitation-sedation dynamics, which are discussed in more detail in this section.

Chapter 2 presents a wavelet analysis of ICU patients' *bivariate* agitation sedation (A-S) time series profiles. New wavelet-based diagnostics based on the concept of a wavelet correlation (WCORR) and wavelet cross-correlation (WCCORR), where the WCCORR is used to determine

lead or lag relationships between bivariate time series on a scale-by-scale basis is also proposed in Chapter 2. These statistics are shown to successfully test the feasibility of the deterministically determined Equations (1.1-1.3) (of Chase et al. (2004)) that provide each patient's simulated A-S series. Chase et al.'s (2004) model serves as a basis for comparison of sedation administration methods, devices, therapeutics, and protocols. Heavy derivative feedback control was also demonstrated by Chase et al. (2004) to be an effective means of managing agitation, given consistent agitation measurement. Implementation of automated feedback infusion controllers based on Chase's type of model can then offer simple and effective drug delivery, without significant increases in drug consumption and expenses.

Specifically in Chapter 2 a suite of proposed wavelet analytic diagnostics, based on estimates of WCORR and WCCORR (from a DWT-MRA and MODWT-MRA analysis), identified 15 ICU patients (patients 2, 4, 7, 9, 10, 11, 12, 21, 22, 27, 28, 32, 33, 34, 35) (out of 37 analysed) for whom their simulated A-S profiles are poor indicators of their true A-S status. The remaining 22 ICU patients analysed were identified via our wavelet indices as "good trackers," that is, their simulated profiles were representative of their true A-S status over the patient's length of stay in ICU. Recall that the simulated profiles are based on the differential equations as derived by Chase et al. (2004) (Table A.2). Thirteen of our wavelets based "poor trackers" were also identified as patients whose simulated A-S profile was a poor fit to their observed A-S rates, by either the kernel smoothing, tracking index and probability band approach of Chase et al. (2004) or by the Average Normalised Density (AND) and Relative Average Normalised Density (RAND) measures of Rudge et al. (2006b). Our wavelet based performance diagnostics are thus shown to be capable of statistically assessing whether an ICU patient's mathematically simulated agitation status over time reflects their true recorded and dynamic (A-S) profile. It is thus established, using wavelets, and in agreement with Chase et al. (2004), Rudge et al. (2006b) and Lee et al. (2005) that these mathematical models and simulations can then be employed as the basis of further improved platforms by which to assess A-S control protocols and procedures.

Chapter 2 builds on the work of Hudson et al. (2004), which was a very preliminary study to assess so-called wavelet signatures for modelling the same ICU agitation-sedation profiles, and to, as in Chapter 2, evaluate "closeness" or "discrimination" of a patient's A-S series, with respect to wavelet scales. To the best of our knowledge, the work in Chapter 2 represents the first application of the DWT and of correlation and cross-correlation based wavelets analysis of agitation-sedation dynamics. A fundamental reason why the DWT approach is used as a tool in Chapter 2 for analysis of the ICU data is that this approach effectively allows for the construction of a time series perfectly from its resultant DWT coefficients. The DWT allows one to partition a time series into pieces that are associated with different scales and time. Wavelets are also useful

for nonparametric problems, as they form sparse representations of functions, including those with discontinuities or other forms of in-homogeneity, this motivates the work in Chapter3.

In Chapter 3, a nonparametric density estimation approach via wavelet smoothing is developed for assessing the validity of deterministic A-S dynamic models against empirical data. A wavelet smooth regression and its density estimation approach, is shown to yield excellent visual graphical assessment tools, in addition to numerical wavelet based metrics, by which to statistically assess the compatibility between an ICU patient's simulated (deterministically modelled) and their true observed (recorded) A-S time series data. The work in Chapter 3 constructs a novel wavelet probability band (WPB) and yields accurate graphical assessments and numerical metrics of compatibility between the modelled and the raw data, at a patient by patient level. In Chapter 3 the density is estimated using wavelet shrinkage methods, based on Bayesian methods. Specifically the minimax estimator is used to obtain a patient specific wavelet tracking coverage index (WTCI). All values of the WTCI are obtained using Bayesian wavelet thresholding, and shown to well reflect poor versus good tracking. A Bayesian approach is suggested in Chapter 3 by which to assess a parametric A-S model – this by constructing a wavelet probability band (WPB) for the proposed model and then checking whether the *nonparametric* regression curve lies within the band. The wavelet probability band (WPB) is shown in Chapter 3 to provide a useful tool to measure the compatibility of the patient's simulated and recorded time series profiles. Moreover, the density profile is then successfully used to define and compute two numerical measures, namely the average normalized wavelet density (ANWD) and relative average normalized wavelet density (RANWD); measures of agreement between the recorded infusion rate and simulated infusion rate. Our WPB method is also shown to be a good tool for detecting regions where the simulated infusion rate (model) performs poorly, thus providing ways to help improve and distil the deterministic A-S model. A so-called Wavelet Time Coverage Index (WTCI) is also developed (see Table 3.2), analogous to the metrics based on a kernel based probability band of Rudge et al. (2006b) and of Chase et al. (2004).

The wavelet regression diagnostics in Chapter 3 (WTCI, ANWD, RANWD, and WPB) identify 13 ICU patients (patients 2, 4, 7, 9, 10, 11, 14, 22, 23, 27, 28, 32 and 33) (out of 37 analysed) for whom their simulated A-S profiles are poor indicators of their true A-S status. Eleven of these 13 "poor trackers" are also identified as poor trackers by the wavelet DWT, wavelet correlation and cross-correlation measures derived in Chapter 2. The WTCI and 90% WPB derived in Chapter 3 thus give strong support for the work in Chapter 2 and vice versa. Seven of these 13 so-called "poor trackers" were also identified by either the kernel smoothing, tracking index and probability

band approach of Chase et al. (2004) or the Relative Average Normalised Density (RAND) measures of Rudge et al. (2006b).

The wavelet methods in Chapter 2 and Chapter 3 confirm that the mathematical, dynamic A-S models of recent studies (Chase et al., 2004; Rudge et al., 2005,2006b; Lee et al., 2005) are suitable for developing more advanced optimal infusion controllers. We conclude that wavelets based modelling can form an appropriate feedback for comparison of improved sedation administration controllers and gain. The simulated nurse's assessment results can then be used as the benchmark for assessing the effects of simple changes to the infusion control system. This study clearly shows that wavelet modelling captures the fundamental dynamics of the agitation-sedation system. Wavelets are thus suitable for clinical implementation.

As far as the author is aware this is the first study to develop a nonparametric wavelet approach to create a wavelet probability band, and associated metrics (RANWD and ANWD) for assessing the validity of deterministic dynamic A-S models against empirical data, based on wavelet smooth regression and wavelet density estimation. Equally important are the resultant visual graphical assessment tools, in addition to the wavelet based numerical metrics for compatibility between the modelled and observed A-S data (see Table A.1 and Table 1.3).

7.3 Brain image analytic methods: Denoising and modelling personality

The work in Chapter 4 investigates single photon emission computed tomography (SPECT) brain image denoising using several wavelet methods (see Figure 4.14). Different methodologies for noise reduction, mainly in the wavelet domain, are compared with a view to gaining insight into which algorithm should optimally be used to find the most reliable estimate of the original image, given its degraded version. Specifically different wavelets based brain image denoising techniques are compared - for two different wavelet transform schemes, namely the DWT and MODWT, used with four different wavelet bases. The aim of the selected designs is to study the suitability of different wavelet bases and also investigate the effect of different window sizes on denoising of 2D images. The *SureShrink* denoising method is shown to be superior. Also the LaDaub8 wavelet base performs well in our image denoising application. Depending on the type of images and the scale of the objects in them, one may, however, prefer different wavelet transforms.

In Chapter 4 the wavelet denoising methods are also compared to independent components analysis (ICA). In the application of Chapter 4, denoising of images is restricted to 2D. However, single photon emission computed tomography (SPECT) brain image data is 3D, and denoising is

generally performed in 3D space to take advantage of the resultant separation of noise and signal in higher dimensions and the availability of volumetric features available in true 3D datasets. Unlike wavelet based denoising methods, ICA based methods use a transform which is estimated from the available data. Therefore, ICA based methods may perform better than wavelet based methods in some applications, this is yet to be rigorously tested.

In Chapter 5 independent component analysis is used to denoise SPECT 3D brain image data (see Figure 1.1 which illustrates the links between chapters of this thesis). Specifically Chapter 5 presents an independent components analysis using FastICA (Hyvärinen, 1998), combined with statistical parametric mapping (SPM) (Friston et al., 2003) as a novel approach by which to model brain function with respect to personality (see Table 5.2). ICA is used principally to denoise the brain image data, prior to modelling brain activations or deactivations, with respect to the psychometric constructs of personality and character (Cloninger, 1994, 2003). The neuroinformatics study in Chapter 5 extends the work of Turner et al. (2003) by using ICA followed by SPM to assess differences in cerebral blood flow (rCBF) between the seven temperament and character axes, the so-called personality traits (Cloninger, 2002, 2008).

The value of using a personality quartile grouping in SPM with t-contrasts to model personality as a predictor of rCBF at a voxel by voxel level is demonstrated in Chapter 5. Indeed the results in Chapter 5 show that a quartile modelling approach allows for the detection of *non-linear* relationships between brain function and the four-quartile levels of the personality traits. These could not have been detected via conventional GLM based linear modelling. Also independent component analysis (ICA) is shown in this study to be an effective method for removing artifacts and separating sources of the brain signals. ICA finds an unmixing matrix that decomposes the SPECT brain image data into a sum of spatially independent components. The SPECT brain image is then reconstructed via the inverse unmixing matrix and ICA. The inverse unmixing matrix represents the relative rCBF strength of the brain regions recruited by the corresponding component map across patients.

The results in Chapter 5 confirm a significant link between regional cerebral blood flow and the Cloninger personality model and confirm and extend the results of Turner et al. (2003). Specifically, this study shows an activation of novelty seeking in the precentral and post central gyri, which corresponds to the motor cortex and the somatic sensory cortex, with deactivations are found in the temporal gyri, occipital lobe and precuneus (Chapter 5). Higher levels of harm avoidance are correlated with significant activations in the limbic lobe and temporal lobe. Individuals who are high in reward dependence are approval seeking, whereas those who are low in reward dependence are aloof. Reward dependence is expected to be related to the circuit of

Papez (Cloninger, 2002), which loops from the hippocampal formation, through the fornix, mammillary body, anterior thalamic nucleus, cingulate gyrus, part of the parahippocampal gyrus and back to the hippocampal formation (Nolte, 1981). This circuit involves the thalamo-cingulate subdivision of the limbic system. Suguira et al. (2000) also showed a significant relationship between reward dependence and decreased blood flow in the paralimbic regions. Likewise, Turner et al. (2003) showed increased reward dependence is related to activations in the frontal and temporal lobes and deactivations in the temporal and occipital lobes. An activation was found in the left middle frontal gyrus in the same region as a negative correlation reported by Suguira et al. (2000).

Persistence exhibited activations in the occipital, temporal, and frontal lobes, but these are not statistically significant (Chapter 5). Deactivations were found in the temporal, frontal, posterior, limbic lobes, and insula (Turner et al., (2003)) also showed activations in the temporal, parietal, occipital and limbic lobes. Deactivations were found in the parietal, temporal and frontal lobes as well as operculum and insula by Turner et al. (2003).

The character traits are expected to be related to higher cognitive functions (Cloninger, 2002) . The brain regions thus expected to be related to the character traits are the thalamo-neocotical system and the prefrontal cortex (Cloninger, 2002). Self-directedness showed activations in the anterior, frontal lobes, and insula. Turner et al. (2003) also showed activations in the frontal lobe only, with deactivations in the precentral gyrus, frontal lobe, temporal lobe and occipital lobe. Cooperativeness was correlated with activations in the temporal, limbic, occipital, frontal, and parietal lobes. Deactivations were found in the sub-lobar (extra-nuclear corpus callosum) and parietal lobe (Chapter 5). Turner et al. (2003) also showed that cooperativeness was associated with the most brain region clusters and specifically with activations in the frontal lobe, limbic lobe and sub cortical gray nuclei. Self transcendence was correlated with activations in the parietal lobes and deactivations in the temporal lobe (Chapter 5). Turner et al. (2003) found activations in the occipital lobe and optic radiation and deactivations in the temporal and parietal lobe. The occipital lobe was found in the activation cluster, but was not significant in this study.

These results derived from our SPM-ICA approach support work showing a biological basis for the TCI model (Cloninger, 2002, 2008) and give credence to a growing body of thought for the need for non-linear models in psychometric research (Cloninger, 2008). Also our research has the potential to provide constructs for drug trial evaluation of depressives (non-normal patients) and help increase the momentum for patient specific drug design and treatment. Recently Turner, Hudson & Joyce (In Prep) established gender specific and non-linear relationships between symptoms of depression and subsets of all seven TCI temperament and character constructs. This

provides support for our conjecture that the relationships between brain function and TCI may be also gender specific and clearly also non-linear as established by our SPM-ICA modelling and also earlier by Turner et al. (2003), which was the first study, to that date, to study all seven of Cloninger's trait. Furthermore, this study (Chapter 5) via our SPM-ICA approach, finds more activation or deactivation clusters in the brain, as related to specific character and temperament traits, than did the study of Turner et al. (2003).

7.4 Detecting change points

The good time-frequency localization properties of wavelets provide a natural motivation for using wavelet methodologies in change point problems, where the main goal is to detect change points and their locations. In Chapter 6 we have developed a new DWT based methodology for change point analysis that essentially uses shifting of the largest wavelet coefficient as a data driven nonparametric procedure for detecting a change point. Specifically, this novel wavelets based change point approach uses modified maximal overlap discrete wavelet transform (MODWT) coefficients to link to a shifting DWT (SDWT) methodology. Our SDWT is thus a combination of the DWT and MODWT for the change point detection problem. The SDWT is compared, via a simulation data set, to some of the existing change point detection methods, namely, Atheoretical Regression Trees (ARTs) and Bai and Perron (Bai & Pearson, 2003, 1998) BP' s method and also to conventional wavelet methods (Donoho & Johnson, 1998; Ogden, 1997).

The shifted DWT (SDWT) based methodology is shown in Chapter 6 to perform accurately because it efficiently portions the time-frequency plane by using short basis functions for high-frequency oscillations and long basis functions for low frequency oscillations. These characteristics of the SDWT always gives one change point for single change point series, which is accurate for long series. SDWT also has the advantage of simplicity since it can be easily implemented with packages that take a very short time to run computationally. Although ART and BP have advantages, when the series are short ART is very prone to over-fitting, with this tendency gradually disappearing as the series becomes longer. On the other hand BP tends to under-fit for shorter series and noisy series. The amount of computation time required for BP, makes it an impractical tool for analysing long time series. The simulations in Chapter 6 have also have provided ways to overcome problems of existing change point detection methods like ART and BP. As far as the author is aware this is the first to date development of a shifted DWT method by which to detect a change point.

SWDT also has an added advantage in that it can also be easily generalized to find multiple change points. This is achieved by way of a binary segmentation procedure, by iterating the process in the two subsequences that surround the first detected change point. Then, if a SWDT

variance based change is detected subsequently, one can split again and iterate the process until no more change points are found in any of the subsequences.

7.5 Future Work

There are a number of areas that can be further investigated to build on the work presented in this thesis. One overarching area is to extend the wavelet transform analysis to allow for more complicated models.

In this thesis modelling brain activations or deactivations, with respect to psychometric constructs of personality and character (Cloninger, 1994, 2008) is achieved using ICA and statistical parametric mapping (SPM) (Friston et al., 2007). An area for further work is to use ICA and wavelets on the psychometric and brain image data, specifically on the low-resolution SPECT images, so as to improve resolution to get the desired clear image. Currently available wavelets based methods for analysing the brain images are mostly developed for the analysis of functional magnetic resonance images (fMRI). Wavelet modelling of cerebral blood flow in relationship to personality, is a clear topic of future work. This could not be achieved in the present study due to computational and time limitations.

One further future study includes parameterizing the wavelet-smoothed time series to identify multiple change points. We foresee the ability to use different types of time series in this wavelet method of analysis. Specifically of interest would be to investigate how best to locate a threshold that finds sudden changes, so as to forecast the time series. The choice of a threshold is a very fundamental issue in wavelet smoothing and there are many of possibilities for choice of the threshold value. It is anticipated that applying general extreme value theory may be one way to achieve this.

Future work will also focus on the validation of the wavelet analytic methods developed here on different data applications, for example:

1. Plant phenomics data, with a view to wavelet modelling of plant health measured in many dimensions versus salinity and other environmental factors. This is part of a Premier Science Research Fund (2010-2012) grant of Lun, Tester, van den Hengel and Hudson.[Ψ]
2. Sleep research data, to statistically model Australia wide railway drivers' (RDs) transition probabilities to sleep and wake (these are the so-called RD-specific (λ, Q) profiles) in relation to time, current, retrospective and prospective anticipatory shift/break, next duty,

[Ψ] Premier Science and Research Fund (PSRF) (2010-2012) "*Plant Image Analysis for the development of stress-tolerant Crops*," Four Chief Investigators: Lun D (UniSA math), Tester M (University of Adelaide, ACPFG), van den Hengel A (University of Adelaide), Hudson IL (UniSA math/stat)

break onset, hours to the next duty onset times. This research aims to extend our wavelets based research to benefit end-users by the creation of new work-related fatigue models, and has clear work, health and safety implications for the railway industry. This is part of a Cooperative Research Centre, Rail Innovation[π] grant (2010-2012) of Ferguson, Darwent, Hudson, Blundell, Rajaratnam and di Milia.

3. Phenological and climate change data – expanding on the work of Hudson et al (2010a) to eight eucalypt species instead of four, to find climatic and species specific wavelet signatures of flowering for southern hemisphere eucalypts (see also Hudson 2010).

[π] CRC Rail Innovation Grant: *Next Generation Fatigue Models:* (2010-2012) *Enhancing the sleep/wake/work database that underpins fatigue-modelling algorithms.* Project Leader Dr Sally Ferguson (Centre of Sleep Research, UniSA); with investigators: Drs David Darwent and Sarah Blunden (Centre of Sleep Research, UniSA), A/Prof Irene Hudson (School of Mathematics and Statistics, UniSA), (A/Prof S Rajaratnam) (School of Psychology, Psychiatry and Psychological Medicine, Monash University) and Dr L di Milia (School of Management and Information Systems, University of Central Queensland)

Appendix A

TableA.1 Overview of Studies on ICU data

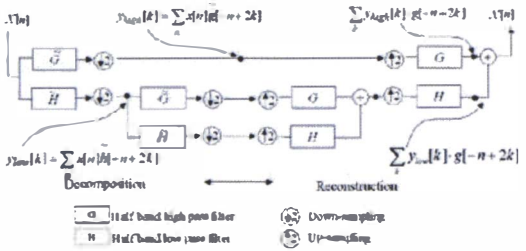
Authors	Equations or Model used	Mathematical methods used	Aims of the study and the performance indicators derived
Kang et al. (Chapter 2)	See equations in the Chase et al. (2004) row below and see the schema of our approach of Chapter 2 below: **DWT Procedure (second stage)**  **DWT analysis and synthesis equations** $X=[X_1, X_2,\ldots, X_N]$, $N=2^J$, DWT analysis equation $W=\mathcal{W}X$, $W=$ discrete wavelet coefficients, $\mathcal{W}=N \times N$ orthonormal matrix $W=\mathcal{W}X, W=\left[W_1 W_2,\cdots,W_J,V_J\right]^T, \mathcal{C}=[\mathcal{W}_1\mathcal{C}_2,\ldots,\mathcal{W}_J,\mathcal{V}_J]$ $X=\mathcal{W}^T W=[\mathcal{C}_1,\mathcal{W}_2,\ldots,\mathcal{C}_J,\mathcal{V}_J]^T[W_1,W_2,\ldots,W_J,V_J]^T$ $=\sum_{j=1}^{J}\mathcal{W}_j^T W_j+\mathcal{V}_J^T V_J \Rightarrow$ DWT synthesis equation **DWT-MRA** $X=\sum_{j=1}^{J}\mathcal{W}_j^T W_j+\mathcal{V}_J^T V_J=\sum_{j=1}^{J} D_j+S_J \Rightarrow$ Additive decomposition(=MRA) $D_j=\mathcal{W}_j^T W_j$: Portion of synthesis due to scale λ_j, jth 'detail' $S_J=\mathcal{V}_J^T V$: 'smooth' of Jth order	Maximal Overlap Discrete Wavelet (MODWT)(Percival & Walden, 2000) Multiresolution analysis (MRA) (Percival & Walden, 2000) DWT-MRA, MODWT-MRA Wavelet shrinkage (Gencay, Selcuk,, & Whitcher, 2002; Ogden, 1997)	Develop a wavelet regression, and wavelet correlation (WCORR) and wavelet cross-correlation (WCCORR) approach for assessing the validity of the deterministic dynamic models against the empirical agitation-sedation data of 37 ICU patients. Provide graphical assessment tools and wavelet based numerical metrics of compatibility between the model and the data via the discrete wavelet transform (DWT), partial DWT (PDWT), maximal overlap DWT (MODWT) and via Multiresolution analysis (MRA). Investigate the lag/lead relationship between the simulated and recorded infusion series on a scale by scale basis via wavelet cross-correlation (WCCORR). Develop performance measures as follows: 1. Modulus of the wavelet correlation at wavelet scale 1, λ_1. 2. Count number (out of 8) of non-significant wavelet correlations at scales λ_j(j=1,2,…,8). 3. Median and 95% CI of the first 5 wavelet correlations at scales λ_j (j=1,2,…,5). Test poor versus good fit or tracking groups via the: Kruskal Wallis test of the indicators in 1, 2, and 3, above for the poor versus good trackers groups.

Table A.1 continued

Kang et al. (Chapter 3)	See equations in the Chase et al. (2004) row below and see the schema of the approach developed in Chapter 3 below: **Wavelet shrinkage (threshold) procedure** Data X_t → DWT → Threshold (minimax) → IDWT → $f(x) \Rightarrow \tilde{f}_J$ (wavelet estimate) $f(x) = \sum_k c_{j_0,k}\phi_{j_0,k}(x) + \sum_{J>J_0}\sum_k d_{J,k}\psi_{J,k}(x)$: a square intergrable density function ϕ : orthogonal scaling function, ψ : mother wavelet $\hat{f}_J(x) = \sum_k \hat{c}_{j_0,k}\phi_{j_0,k}(x) + \sum_{J>J_0}\sum_k \hat{d}_{J,k}\psi_{J,k}(x)$: wavelet estimator for $f(x)$ at J level $\hat{c}_{J,k} = \frac{1}{n}\sum_{t=1}^{n}\phi_{J,k}(X_t),\ \hat{d}_{J,k} = \frac{1}{n}\sum_{i=1}^{n}\psi_{J,k}(X_t)$	Density estimation (Ogden, 1997; Nason, 2008) Wavelet thresholding via BayesThresh methods (Abramovich & Silverman, 1998; Nason, 2008) Wavelet shrinkage (threshold)(Gencay, Selcuk & Whitcher, 2002; Ogden, 1997; Nason, 2008)	Develop a density estimation approach via wavelet smoothing for assessing the validity of deterministic dynamic models (simulated profiles) against the empirical / recorded data. Construct a wavelet probability band (WPB). Provide graphical assessment and numerical metrics of compatibility between the model and the recorded agitation-sedation data. Develop performance measures as follows: 1. Average normalized wavelet density (ANWD). 2. Relative average normalized wavelet density (RANWD). 3. Median of the Wavelet Time Coverage Index (WTCI).
Rudge et al. (2006 a, 2006b)	I. Pharmocokinetics of morphine $V_c^o \frac{dC_c^o}{dt} = -\left(K_{CL}^o + K_{ce}^o + K_{cp}^o\right)C_c^o + P^oU + K_{ce}^o C_e^o + K_{pc}^o C_p^o$ $V_p^o \frac{dC_p^o}{dt} = -K_{pc}^o C_p^o + K_{cp}^o C_c^o,\ V_e^o \frac{dC_e^o}{dt} = -K_{ec}^o C_e^o + K_{ce}^o C_c^o$ II. Pharmocokinetics of midazolam $V_c^z \frac{dC_c^z}{dt} = -\left(K_{CL}^z + K_{ce}^z\right)C_c^z + P^zU + K_{ec}^z C_e^z,\ V_e^z \frac{dC_e^z}{dt} = -K_{ec}^z C_e^z + K_{ce}^z C_c^z$ III. Pharmacodynamics of morphine and midazolam $\frac{dA}{dt} = w_1 S - w_2 K_T \int_0^t E_{comb}(\zeta) e^{-K_T(t-\zeta)} d\zeta,\ E_{Comb} = E_0 + \left[E_{max}(\theta) - E_0\right] - \frac{\left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}{1 + \left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}$	Kernel smoothing :Chebychev's inequality for probability band(Chase et al., 2004) Relative average normalised density (RAND) Average normalised density (AND)	Develop a physiologically representative model that incorporates endogenous agitation reduction (EAR). Use performance measures as follows: 1. RTD: relative total dose (RTD) expresses the total dose administered in the simulation as a percentage of the actual total recorded dose. 2. Relative average normalised density (RAND) measures how probabilistically similar the model outputs are to the smoothed data, and hence the degree of compatibility between the model and the empirical data. 3. Percentage time in band (TIB)

Table A.1 continued

Rudge et al (2005)	The agitation-sedation system model: Phamarcokinetic model (Wood & Wood, 1990) adding patient agitation as a third state variable $\frac{dC_e}{dt} = -K_1 C_e + \frac{U}{V_d}$ $\frac{dC_p}{dt} = -K_2 C_p + K_3 C_e$ $\frac{dA}{dt} = w_1 S - w_2 \int_0^t C_p(\tau) e^{-K_4(t-\tau)} d\tau$	Infinite Impluse Response (IIR) filter (Rorabaugh, 1998) Proportional-Derivative (PD) control with agitation for infusion rate (U) Moving blocks bootstrap (Efron & Tibshirani, 1993) Tracking Index (TI)	Develop a control model to capture the essential dynamics of the agitation-sedation system. Use performance measures as follows: 1. $U = K_p A + K_d A$ for the infusion rate. 2. Tracking Index (TI): Quantitative parameter to indicate how well the simulated infusion rate profile represents the average recorded infusion profile over the entire time series.
Lee et al. (2005)	I. Pharmocokinetics of morphine $V_c^o \frac{dC_c^o}{dt} = -\left(K_{CL}^o + K_{ce}^o + K_{cp}^o\right) C_c^o + P^o U + K_{ec}^o C_e^o + K_{pc}^o C_p^o$ $V_p^o \frac{dC_p^o}{dt} = -K_{pc}^o C_p^o + K_{cp}^o C_c^o, \quad V_e^o \frac{dC_e^o}{dt} = -K_{ec}^o C_e^o + K_{ce}^o C_c^o$ II. Pharmocokinetics of midazolam $V_c^s \frac{dC_c^s}{dt} = -\left(K_{CL}^s + K_{ce}^s\right) C_c^s + P^s U + K_{ec}^s C_e^s, \quad V_e^s \frac{dC_e^s}{dt} = -K_{ec}^s C_e^s + K_{ce}^s C_c^s$ III. Pharmacodynamics of morphine and midazolam $\frac{dA}{dt} = w_1 S - w_2 K_T \int_0^t E_{comb}(\zeta) e^{-K_T(t-\zeta)} d\zeta, \quad E_{Comb} = E_0 + \left[E_{max}(\theta) - E_0\right] - \frac{\left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}{1 + \left(\frac{C_o + C_S}{C_{50}(\theta)}\right)^{\gamma(\theta)}}$	Kernel regression(Wand & Jones, 1995) Kernel density estimation: marginal density function of the regression function estimate Nonparametric regression: Chebychev's inequality for probability band(Chase et al., 2004)	Develop a nonparametric approach for assessing the validity of deterministic dynamics models against empirical data. Use performance measures as follows: 1. Kernel regression and kernel density estimation to yield visual graphical assess the data. 2. Construct a probability band for the nonparametric regression curve and check whether the proposed model lies within the band. 3. 3. Average normalised density (AND) to measure how well the simulated values coincide with the maximum density at every time point and relative average normalised density (RAND).

Table A.1 continued

Chase et al (2004)	The agitation-sedation system model: Phamarcokinetic model (Wood & Wood, 1990) adding patient agitation as a third state variable $\frac{dC_c}{dt} = -K_1 C_c + \frac{U}{V_d}$ $\frac{dC_p}{dt} = -K_2 C_p + K_3 C_c$ $\frac{dA}{dt} = w_1 S - w_2 \int_0^t C_p(\tau) e^{-K_4(t-\tau)} d\tau$ Uniform kernel with bandwidwidth h(Wand & Jones, 1995) $K_h = \begin{cases} 0 & \text{if } t < -\frac{h}{2} \\ 1 & \text{if } t - \frac{h}{2} < t \le \frac{h}{2} \\ 0 & \text{if } t > \frac{h}{2} \end{cases}$	Infinite Impluse Response (IIR) filter (Rorabaugh, 1998) Proportional-Derivative (PD) control with agitation for infusion rate (U) Tracking Index (TI) Chebychev's inequality for probability band (Chase et al., 2004)	Develop a mathematical model to capture the essential dynamics of the agitation-sedation system and develop statistical validity using the recorded infusion data for 37 ICU patients. Use performance measures as follows 1. Kernel smoothing using the uniform kerne.l. 2. Tracking Index (TI). 3. Moving blocks bootstrap to gain an understanding of the reliability of the TI for a given patient's infusion profile 4. 90% Probability Band by definition the range within at least 90% of the time, the estimated mean value of the recorded infusion rate lies within the band.

C_c, C_p and C_e are, respectively, the drug concentrations (mg L^{-1}) in the central, peripheral and effect compartments;, U is the intravenous infusion rate; V_d , V_c, V_p and V_e, respectively, the volume of distribution, the distribution volumes (L) of the central, peripheral and effect compartments; A is an agitation index, S is the stimulus invoking agitation; K_1–K_3 are parameters relating to drug elimination and transport and K_{ij} the transfer rate (L min^{-1}) from compartment i to compartment j; K_{CL} the drug clearance (L min^{-1}); K_T the effect, and w_1 and w_2 are relative weighting coefficients of the stimulus and drug effect, respectively. Time is represented by t, and τ is the variable of integration in the convolution integral V_c, V_p and V_e, respectively, the distribution volumes (L) of the central, peripheral and effect compartments; U the intravenous infusion rate (mL min^{-1}); A an agitation index; S the stimulus invoking agitation; K_{ij} the transfer rate (L min^{-1}) from compartment i to compartment j; K_{CL} the drug clearance (L min^{-1}); K_T the effect time constant (min^{-1}); P^o and P^s are the concentrations of morphine and midazolam, respectively (mgmL^{-1}), where terms with superscript 'o' relate to the opioid morphine, and terms with superscript 's' relate to the sedative midazolam. Time is represented by t (min), the variable of integration, and the terms w_1 and w_2 are the relative weights of stimulus and cumulative effect, representing the patient sensitivity. Finally, Ecomb is the combined

pharmacodynamic effect of the individual effect site drug concentrations of morphine and midazolam determined using response surface modeling as defined in Minto. et al (2000).

Table A.2 Overview of Studies on Baseline CBF and Personality Traits (sourced directly from Hermes (2007) with permission)

Authors	Participants	Personality trait	Personality trait measure	Imaging technique	Regions analyzed	Significant associations found
Carlsson et al. (2000)	24 men (M = 23 y)	Anxiety (A) Creativity (C)	STAI CFT	133-xenon	Frontal, temporal, parietal, and occipital areas (mean relCBF), hemispheres (mean CBF)	A: no significant association to hemispheric CBF and prefrontal relCBF C: greater hemispheric (left and right) CBF in highly vs. low creative subjects
Ebmeier et al. (1994)	12 men, 21 women (M = 54 y)	Extraversion/ Introversion (E/I) Neuroticism (N) Psychoticism (P)	EPQ	SPECT	Frontal, temporal, and parietal areas, basal ganglia, thalamus, ACC (mean relCBF acquired on 2 slices)	E/I: ACC N: no significant associations P: no significant associations
Johnson et al. (1999)	10 men, 8 women (M = 29 y)	Extraversion/ Introversion (E/I)	NEO-FFI	PET	Whole brain (voxel-based analysis)	E: temporal lobes, ACC, posterior thalamus, right posterior insula, left amygdala I: frontal areas, right anterior temporal cortex, anterior thalamus, left hippocampus, right anterior insula/ putamen, left MCC
Mathew et al. (1984)	51 women (M = 32 y)	Extraversion/ Introversion (E/I) Neuroticism (N)	EPI	133-xenon	Frontal, temporal, parietal, and occipital areas, hemispheres (mean CBF)	E/I: all areas N: no significant associations
O'Gorman et al. (2006)	15 men, 15 women (M = 28 y)	Extraversion/ Introversion (E/I) Neuroticism (N) Psychoticism (P) Novelty seeking (NS) Harm avoidance (HA) Persistence (PS)	EPQ-R TCI	CASL	Whole brain (voxel-based analysis)	E/I: basal ganglia, thalamus, inferior frontal gyrus, cerebellum, cuneus N: no significant associations P: right thalamus, right basal ganglia NS: cerebellum, left thalamus, cuneus HA: cerebellum, cuneus, medial frontal gyrus PS: basal ganglia
Stenberg et al. (1993)	8 men, 9 women (M = 29 y)	Extraversion/ Introversion (E/I) Anxiety (Anx) Impulsivity (IM)	EPI KSP	133-xenon	Frontal, temporal, parietal, and occipital areas (mean relCBF), global CBF (mean CBF)	E/I: greater temporal relCBF for I than for E; significant association between I and global CBF in women Anx: significant association between temporal relCBF and anxiety IM: no significant associations
Stenberg et al. (1993)	8 men, 9 women (M = 29 y)	Extraversion/ Introversion (E/I) Anxiety (Anx) Impulsivity (IM)	EPI KSP	133-xenon	Frontal, temporal, parietal, and occipital areas (mean relCBF), global CBF (mean CBF)	E/I: greater temporal relCBF for I than for E; significant association between I and global CBF in women Anx: significant association between temporal relCBF and anxiety IM: no significant associations

Table A.2 (continued)

Sugiura et al. (2000)	13 men, 17 women (range 26-62 y)	Novelty seeking (NS) Harm avoidance (HA) Reward dependence (RD)	TCI[a]	SPECT	Whole cortex (voxel-based analysis)	NS: left ACC, insula HA: several frontal, temporal, parietal, and paralimbic areas RD: several frontal, temporal, and paralimbic areas
Tankard et al. (2003)	30 men (M = 68 y)	Anxiety	STPI	SPECT	Prefrontal areas, hemispheres (mean relCBF)	No significant associations
Turner et al. (2003)	20 men (range 20-33)	Novelty seeking (NS) Harm avoidance (HA) Reward dependence (RD) Persistence (PS) Self directedness (SD) Cooperativeness (CO) Self transcendence (ST)	TCI	SPECT	Whole brain (voxel-based analysis)	Several associations are reported for each of the analyzed traits.
Zald et al. (2002)[b]	Sample 1: 28 men, 23 women (range 18-50 y); Sample 2: 24 men, 14 women (range 19-55 y)	Negative affect	PANAS	PET	Whole brain (voxel-based analysis)	Sample 1: ventromedial prefrontal cortex, left parainsular region Sample 2: ventromedial prefrontal cortex

Note. Cerebral blood flow (CBF) was always recorded in a resting situation and was reported in absolute values or relative to a reference value (relCBF). CBF was either measured with the 133-xenon inhalation technique, single-photon emission computed tomography (SPECT), positron emission tomography (PET), or with continuous arterial spin labeling (CASL).

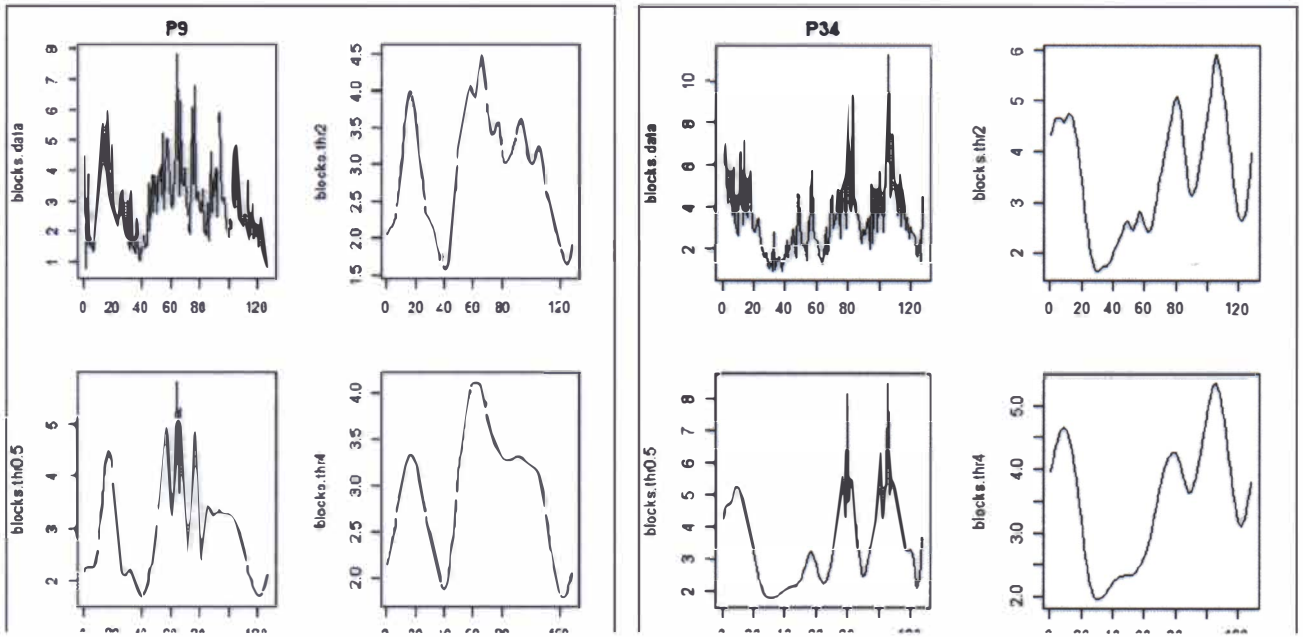
The personality traits were assessed with the Spielberger state-trait anxiety inventory (STAI, Spielberger, 1983), Creative Functioning Test (CFT, Smith & Carlsson, 1990), Eysenck Personality Questionnaire (EPQ, Eysenck & Eysenck, 1975), NEO Five-Factor Inventory (NEO-FFI, Costa & McCrae, 1992), Eysenck Personality Inventory (EPI, Eysenck, 1968), Revised Eysenck Personality Questionnaire (EPQ-R, Eysenck & Eysenck, 1991), Temperament and Character Inventory (TCI, Cloninger, Przybeck, Svrakic, & Wetzel, 1994), Karolinska Scales of Personality (KSP, Schalling, Edman, & Asberg, 1983), State-Trait Personality Inventory (STPI, Spielberger, 1979), and Positive and Negative Affect Scales (PANAS, Watson, Clark, & Tellegen, 1988).

M=Mean age, y = years, ACC = anterior cingulate cortex, MCC = medial cingulate cortex.

[a] A Japanese version of the TCI was used (Kijima et al., 1996).

Table A.3 37 ICU data

Patient No.	Data
1	3601
2	6421
3	6541
4	4921
5	2941
6	5701
7	3901
8	10561
9	8581
10	20701
11	6721
12	8521
13	5161
14	3001
15	4981
16	13621
17	5941
18	4681
19	7921
20	9661
21	3721
22	9661
23	3481
24	8461
25	3841
26	3901
27	13441
28	12241
29	3241
30	3661
31	18301
32	15181
33	25261
34	8101
35	12721
36	3481
37	7501



FigureA.1 Varying the values of the hyperparameters α and β=1.

TableA.4 WPB values by varying the hyperparameters α and β=1.

WPB	α= 0, β=1	α= 0.5, β=1	α= 2, β=1	α= 4, β=1
P9	87.5 %	57.8%	46.8 %	34.5%
P34	75%	48.4%	31.25%	20.3%

Appendix B

Table B.1 The temperament and Character Index traits, 25 subscales and descriptions.

Trait	Definition	Subscale	Definition/Description
NS	Novelty Seeking	NS1	Exploratory Excitability vs Stoic
		NS2	Impulsiveness vs Reflection
		NS3	Extravagance vs Reserve
		NS4	Disorderliness vs Regimentation
HA	Harm Avoidance	HA1	Anticipatory Worry vs Uninhibited Optimism
		HA2	Fear of Uncertainty vs Confidence
		HA3	Shyness with Strangers vs Gregariousness
		HA4	Fatigability & asthenia vs Vigor
RD	Reward Dependence	RD1	Sentimentality vs Insensitivity
		RD3	Attachment vs Detachment
		RD4	Dependence vs Independence
P	Persistence	RD2	Persistence vs Irresoluteness
S	Self Directedness	S1	Responsibility vs Blaming
		S2	Purposefulness vs Lack of Goal Direction
		S3	Resourcefulness
		S4	Self-Acceptance vs Self striving
		S5	Congruent second Nature
C	Cooperativeness	C1	Social Acceptance vs social Intolerance
		C2	Empathy vs Social Disinterest
		C3	Helpfulness vs Unhelpfulness
		C4	Compassion vs Revengefulness
		C5	Integrated Conscience
ST	Self Transcendence	ST1	Self-Forgetfulness vs Self-Conscious Experience
		ST2	Transpersonal Identification vs Self-Isolation
		ST3	Spiritual Acceptance vs Rational Materialism

Table B.2 Contrast results for harm avoidance, Brodmann's areas=BA

Harm Avoidance	Cluster (x, y, z)	K_E	p-value
Activation Q2Q3			
Left Frontal Lobe: middle frontal gyrus BA8	(-46,26,40)	67	0.028
Left Frontal Lobe: sub-gyral	(-34,-14,38)	42	0.026
Activation Q3Q4			
Right Limbic Lobe: posterior cingulate BA23, *Right Anterior Lobe:* Culmen	(4,-46,22)	920	**0.000**
Right Temporal Lobe: sub-gyral	(56,10,0)	300	**0.000**
Right Limbic Lobe: cingulate gyrus BA24	(8, -2, 16)	203	**0.001**
Right Occipital Lobe: middle occipital gyrus BA19 *Right Temporal Lobe:* middle temporal gyrusBA39	(44,-78,12)	215	**0.001**
Left Occipital Lobe: middle occipital gyrus	(-30,-96,10)	44	0.080
Right Frontal Lobe: sub -gyral	(24,22,14)	63	0.041

Table B.3 Contrast results for reward dependence.

Reward Dependence	Cluster (x, y, z)	K_E	p-value
Activation Q1Q2			
Left Limbic Lobe: cingulated gyrus	(-16,-32,30)	63	0.040
Left Temporal Lobe: superior temporal gyrus	(-36,22,18)	316	0.019
Left Frontal Lobe: sub-gyral	(-36,46,-4)	65	0.020
Left Frontal Lobe: middle frontal gyrus			
Right Parietal Lobe: Sub-gyral,	(26,-32,44)	86	0.069
Right Parietal Lobe: precuneus	(10,-52,58)	63	0.021
Right Frontal Lobe: Precentral gyrus,	(12,-22,64)	60	0.028
Right Limbic Lobe: cingulated gyrus	(16,-40,42)	70	0.030
Deactivation Q1Q2			
Right Limbic Lobe: parahippocampal gyrus	(40,-20,-18)	44	0.476
Right Temporal Lobe: sub-gyral			
Activation Q1Q3			
Right Frontal Lobe: middle frontal gyrus, sub-gyral, superior frontal gyrus	(30,58,18)	316	**0.007**
Left Frontal Lobe: superior frontal gyrus	(-6,18,62)	1269	**0.000**
Left Limbic Lobe: cingulate gyrus BA32			
Left Temporal Lobe: sub-gyral, inferior temporal gyrus BA20	(-44,-44,-6)	298	**0.001**
Right Frontal Lobe: middle frontal gyrus BA6,	(50,12,46)	321	**0.006**
Left Frontal Lobe: sub-gyral, precentral gyrus BA6	(-34,20,14)	304	**0.000**
Left Parietal Lobe: postcentral gyrus BA2	(-54,-26,42)	331	**0.005**
Left Frontal Lobe: middle frontal gyrus BA9			
Right Sub-lobar: insula	(42,24,10)	282	0.015
Right Frontal Lobe: inferior frontal gyrus			

Activation Q1Q4			
Left Frontal Lobe: middle frontal gyrus BA9	(-24,40,36)	684	**0.000**
Left Frontal Lobe: sub-gyral, precentral gyrus	(-36,-16,30)	107	0.008
Left Limbic Lobe: anterior cingulate	(-16,28,0)	454	**0.001**
Left sub-lobar: lentiform nucleus, extra-nuclear, insula	(-26,-12,6)	415	**0.001**
Left Occipital Lobe: lingual gyrus	(-16,-82,2)	126	0.604
Activation Q2Q3			
Left Frontal Lobe: sub-gyral	(-10,-20,60)	263	**0.000**
Left Parietal Lobe: inferior pariental lobule BA40	(-44,-42,56)	158	**0.002**
Left Sub-lobar: insula	(-50,-28,18)	382	**0.005**
Left Occipital Lobe: middle occipital gyrus	(-16,-84,14)	52	0.054
Right Sub-Lobar: thalamus	(12,-18,4)	133	**0.004**
Left Limbic Lobe: cingulate gyrus BA32	(-6,16,32)	76	0.023
Right Occipital Lobe: ligual gyrus	(18,-88,-8)	70	0.028
Left Frontal Lobe: Middle frontal gyrus BA6, precentral gyrus, sub-gyral	(-32,-2,48)	157	**0.002**
Left Frontal Lobe: middial frontal gyrus BA10, superior frontal gyrus	(-2,54,20)	113	**0.007**
Left Sub-Lobar: insula BA13	(-36,16,12)	114	**0.007**
Left Frontal Lobe: precentral syrus	(-60,0,16)	97	0.012
Activation Q2Q4			
Left Frontal Lobe: sub-gyral,	(-34,-10,32)	96	0.015
Right Frontal Lobe: inferior frontal gyrus	(50,38,12)	78	0.026
Right Parietal Lobe: postcentral gyrus	(42,-30,52)	89	0.010
Right Frontal Lobe: sub-gyral			
Left Limbic Lobe: parahippocampal gyrus	(-42,-18,-8)	40	0.097
Deactivation Q2Q4			
Right Pariental Lobe: postcentral gyrus	(42,-30,52)	89	0.019
Left Frontal Lobe: middle frontal gyrus BA8	(-20,36,34)	120	0.008
Left Temporal Lobe: sub-gyral	(-42,-18,-8)	40	0.097
Activation Q3Q4			
Right Limbic Lobe: anterior cingulate	(18,40,6)	96	0.292
Right Frontal Lobe: sub-gyral			

Table B.4 Contrast results for persistence.

Persistence	Cluster (x, y, z)	K_E	p-value
Activation Q1Q3			
Left Occipital Lobe:	(-26,-80,8)	55	0.018
middle occipital gyrus			
Activation Q1Q4			
Left Sub-lobar:	(-2,16,8)	86	0.014
lateral ventricle			
Left Frontal Lobe:	(-24,60,24)	74	0.020
superior frontal gyrus			
Deactivation Q1Q4			
Right Temporal Lobe:	(46,-22,-30)	50	0.015
inferior temporal gyrus BA20			
Right Frontal Lobe:	(28,12,-12)	92	0.018
inferior frontal gyrus			
Right Sub-lobar:			
Insula BA13, extra-nuclear			
Right Posterior Lobe:	(24,-42,-38)	56	0.024
cerebellar tonsil			
Deactivation Q2Q3			
Right Frontal Lobe:	(18,46,-4)	288	0.015
middle frontal gyrus BA10, superior frontal gyrus, sub-gyral			
Right Limbic Lobe:			
Anterior cingulate BA32			
Activation Q3Q4			
Right Frontal Lobe:	(22,-8,58)	80	0.015
middle frontal gyrus BA6			
Left sub-lobar:	(-10,6,16)	254	0.009
extra-nuclear, caudate			

Table B.5 Contrast results for Self Directedness.

Self Directedness	Cluster (x, y, z)	K_E	p-value
Activation Q1Q3			
Right Temporal Lobe:	(50,-22,-6)	44	0.071
middle temporal gyrus			
Right sub-lobar:			
insular BA13			
Left Posterior Lobe:	(-18,-80,-18)	152	0.010
declive			
Left Occipital Lobe:			
middle occipital gyrus BA18			
Left Parietal Lobe:	(-2,-38,64)	85	0.017
paracentral lobule			
Left Frontal Lobe:			
sub-gyral			
Activation Q1Q4			
Left Anterior Lobe:	(-8,-52,-6)	258	**0.000**
culmen			
Right Pariental Lobe:	(50,-66,42)	1897	0.070
inferior pariental lobule BA39			
Right Temporal Lobe:			
Middle temporal gyrus BA37			
Activation Q2Q3			
Left Frontal Lobe:	(-30,-26,34)	305	**0.000**
sub-gyral			
Left Sub-lobar:			
insular, extra-nuclear			
Right Frontal Lobe:	(22,-38,32)	145	**0.002**
sub-gyral			
Right Frontal Lobe:	(60,10,36)	118	**0.005**
middle frontal gyrus BA9			
Right Frontal Lobe:	(18,32,34)	222	**0.000**
middle frontal gyrus BA8, sub-gyral, superior frontal gyrus			
Left Frontal Lobe:	(-2,-34,58)	362	**0.000**
paracentral lobule BA14, paracentral lobule BA4			
Right Frontal Lobe:	(44,-8,24)	41	0.074
precentral gyrus, middle frontal gyrus BA6			
Activation Q2Q4			
Right Frontal Lobe:	(38,56,2)	368	**0.000**
middle frontal gyrus BA10, sub-gyral			
Inferior frontal gyrus			
Left Limbic Lobe:			
cingulate gyrus BA23	(-2,-22,30)	105	0.010
Right Limbic Lobe:			
cingulate gyrus BA24			
Left Frontal Lobe:	(-2,-4,66)	476	**0.000**
superior frontal gyrus, medial frontal gyrus			
Right Pariental Lobe:	(56,-16,32)	44	0.076
postcentral gyrus BA3			

Table B.6 Contrast results for Cooperativeness

Cooperativeness	Cluster (x, y, z)	K_E	p-value
Activation Q1Q2			
Left Limbic Lobe: uncus BA34	(-16,-4,-20)	96	0.383
Left Parietal Lobe: sub-gyral	(-24,-46,24)	41	0.554
Right Limbic Lobe: Parahippocampal BA36	(26,-20,-26)	82	0.302
Right Occipital Lobe: lingual gyrus	(8,-70,2)	96	0.383
Right Limbic Lobe: posterior cinulate			
Right Occipital Lobe: cuneus BA19	(18,-88,30)	104	0.334
Right Limbic Lobe: uncus	(28,0,-22)	64	0.452
Right Limbic Lobe: posterior cinulate BA29	(8,-46,16)	40	0.560
Deactivation Q1Q3			
Left Frontal Lobe: middle frontal gyrus BA9	(-54,18,26)	382	0.048
Left Occipital Lobe: middle occipital gyrus BA18	(-24,-90,20)	762	0.008
Left Posterior Lobe: declive			
Left Temporal Lobe: Middle temporal gyrus BA21, inferior temporal gyrus BA37	(-60,-62,2)	217	0.125
Left Sub-lobar: extra-nuclear corpus callosum	(-12,-38,8)	944	**0.004**
Left Limbic Lobe: cingulate gyrus BA31			
Right Parietal Lobe: postcentral gyrus	(56,-28,44)	1241	**0.001**
Right Sub-lobar: extra-nuclear			
Right Posterior Lobe: cerebellar tonsil	(28,-36,-38)	235	0.112
Left Temporal Lobe: superior temporal gyrus	(-58,-44,20)	114	0.012
Left Parietal Lobe: inferior parietal lobule BA40			
Right Frontal Lobe: sub-gyral	(22,6,28)	67	0.013
Left Frontal Lobe: sub-gyral			
Left Temporal Lobe: middle temporal gyrus	(-42,44,4)	55	0.017
Activation Q1Q4			
Left Frontal Lobe: precentral gyrus	(-64,2,8)	124	0.008
Left Anterior Lobe: culmen	(-6,-48,-2)	130	0.017
Left Limbic Lobe: cingulate gyrus	(-16,-22,36)	64	0.019
Left Parietal Lobe: superior parietal lobule	(-22,-68,56)	54	0.019
Left Occipital Lobe: precuneus	(-24,-72,20)	108	0.025
Left Parietal Lobe: sub-gyral	(-24,-40,44)	45	0.021
Left Temporal Lobe:			

sub-gyral			
Right Occipital Lobe:	(40,-70,-12)	50	0.027
fusiform gyrus, sub-gyral			
Left Frontal Lobe:	(-30,-6,56)	59	0.030
middle frontal gyrus, precentral gyrus BA6			
Deactivation Q2Q3			
Left Temporal Lobe:	(-52,-18,-26)	183	**0.000**
fusiform gyrus, sub-gyral			
Right Occipital Lobe:	(18,-86,22)	69	0.019
cuneus BA18			
Left Limbic Lobe:	(-16,10,36)	97	**0.007**
cingulate gyrus			
Left Frontal Lobe:			
cingulate gyrus			
Left Limbic Lobe:	(-16,-4,-20)	59	0.028
uncus BA34			
Right Occipital Lobe:	(12,-68,2)	119	**0.003**
lingual gyrus BA18			
Left Frontal Lobe:	(-18,42,2)	181	**0.001**
Sub-gyral			
Right sub-lobar:	(16,-12,-6)	72	0.017
extra-nuclear, thalamus			
Right Parietal Lobe:	(38,-54,32)	139	**0.002**
angular gyrus, inferior parietal lobule			
Right Frontal Lobe:	(46,-8,30)	58	0.029
precentral gyrus			
Right Occipital Lobe:	(18,-68,26)	99	**0.006**
precuneus, cuneus			
Left Anterior Lobe:	(-44,-48,-28)	114	**0.004**
culmen			
Left Posterior Lobe:			
cerebellar tonsil			
Left Temporal Lobe:			
fusiform gyrus			
Right Temporal Lobe:	(52,-4,-14)	55	0.033
middle temporal gyrus			
Left Frontal Lobe:	(-28,18,54)	97	**0.004**
middle frontal gyrus BA6, superior frontal gyrus BA8			
Right Frontal Lobe:	(24,6,24)	40	0.063
sub-gyral			
Right Sub-lobar:	(44,-18,-2)	43	0.055
Insula BA13			
Left parietal Lobe:	(-42,-54,50)	50	0.041
Inferior parietal lobule BA40			
Deactivation Q2Q4			
Left Temporal Lobe:	(-42,-6,-22)	75	0.034
sub-gyral			
Activation Q3Q4			
Left Frontal Lobe:	(-26,28,-2)	194	**0.000**
sub-gyral			
Left Limbic Lobe:			
anterior cingulate			
Right Temporal Lobe:	(54,-4,-16)	118	**0.003**
middle temporal gyrus BA21			
Left Occipital Lobe:	(-24,-90,12)	370	**0.000**
middle occipital gyrus			
Left Limbic Lobe:	(-6,-42,30)	225	**0.000**
cingulate gyrus BA31			
Left Sub-lobar:			
lateral ventricle			
Inter-Hemispheric			

Right Frontal Lobe: sub-gyral, superior frontal gyrus BA10 middle frontal gyrus BA9	(24,44,12)	181	**0.000**
Right Anterior Lobe: culmen	(4,-36,-12)	51	0.033
Right Midbrain			
Left Posterior Lobe: tuber	(-48,-48,-24)	88	**0.007**
Left Temporal Lobe: inferior temporal gyrus BA37			
Left Posterior Lobe: declive	(-14,-78,-20)	172	**0.000**
Inter-Hemispheric			
Right Sub-lobar: insula BA47	(28,20,-6)	58	0.024
Right Limbic Lobe: cingulate gyrus BA32	(10,6,42)	121	**0.002**
Right Temporal Lobe: fusiform gyrus	(44,-58,-12)	156	**0.001**
Left Parietal Lobe: inferior parietal lobule BA40	(-52,-58,40)	57	0.025
Right Sub-lobar: thalamus	(14,-18,-4)	103	0.004
Right Midbrain			
Right Occipital Lobe: Lingual gyrus BA18	(14,-68,2)	42	0.005
Right Temporal Lobe: sub-gyral	(22,-68,22)	90	**0.007**
Right Parietal Lobe: precuneus			
Right Frontal Lobe: middle frontal gyrus	(48,34,20)	46	0.041

Table B.7 Contrast results for self-transcendence.

Self Transcendence	Cluster (x, y, z)	K_E	p-value
DeactivationQ1Q2			
Left Temporal Lobe: superior temporal gyrus	(-30,16,-34)	241	**0.000**
Right Anterior Lobe: culmen	(44,-38,-28)	118	0.008
Left Occipital Lobe: sub-gyral	(-28,-54,-2)	110	0.010
Left Sub-lobar: extra-nuclear			
Left Frontal Lobe: inferior frontal gyrus	(-36,34,0)	52	0.061
Right Temporal Lobe: middle temporal gyrus	(42,-54,6)	41	0.091
Deactivation Q1Q4			
Left Frontal Lobe: Precentral gyrus BA4	(-38,-14,50)	55	0.023
Activation Q2Q3			
Left Frontal Lobe: middle frontal gyrus	(-40,56,8)	123	0.013
Left Sub-lobar: lateral ventricle	(-16,-42,14)	801	0.019
Inter-Hemispheric			
Right Limbic Lobe: cingulate gyrus			
Left Frontal Lobe: middle frontal gyrus, sub-gyral	(-20,4,42)	167	0.010
Left Parietal Lobe: sub-gyral	(-46,-14,20)	189	**0.007**
Left Frontal Lobe: Sub-gyral			
Right Frontal lobe: sub-gyral	(42,4,14)		0.026
Right Sub-lobar: Insula			
Activation Q2Q4			
Left Frontal Lobe: superior frontal gyrus BA9	(-14,58,34)	54	0.041
Right Limbic Lobe: cingulate gyrus	(16,-30,36)	52	0.044
Left Frontal Lobe: sub-gyral	(-30,-20,36)	55	0.039
Right Occipital Lobe: Middle occipital gyrus BA18	(30,-94,10)	101	0.008
Deactivation Q3Q4			
Left Sub-lobar: lateral ventericle	(-20,-34,14)	46	0.030

References

Abramovich, F., Bailey, T. C., & Sapatinas, T. (2000). Wavelet analysis and its statistical applications. *The Statistician, 49*(1), 1-29.

Abramovich, F., & Benjamini, Y. (1995). Thresholding of wavelet coefficients as multiple hypotheses testing procedure. *Lecture Notes in Statistics, 103*, 5-14.

Abramovich, F., Besbeas, P., & Sapatinas, T. (2002). Empirical Bayes approach to block wavelet function estimation. *Computational Statistics and Data Analysis, 39*(4), 435-451.

Abramovich, F., Sapatinas, T., & Silverman, B. W. (1998a). Wavelet thresholding via a Bayesian approach. *Journal of the Royal Statistical Society. Series B, Statistical Methodology*, 725-749.

Abramovich, F., & Silverman, B. W. (1998b). Wavelet decomposition approaches to statistical inverse problems. *Biometrika, 85*(1), 115-129.

Abu-Rezq, A. N., Tolba, A. S., Khuwaja, G. A., & Foda, S. G. (1999). Best Parameters Selection for Wavelet Packet-Based Compression of Magnetic Resonance Images. *Computers and Biomedical Research, 32*(5), 449-469.

Achard, S., & Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS computational biology, 3*(2), e17.

Achard, S., Salvador, R., Whitcher, B., Suckling, J., & Bullmore, E. (2006). A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs. *Journal of Neuroscience, 26*(1), 63-72.

Acharyya, R., & Ham, F. M. (2008). A New Approach for Blind Source Separation of Convolutive Sources - Wavelet Based Separation Using Shrinkage Function: VDM Verlag.

Acharyya, R. (2008). *A New Approach for Blind Source Separation of Convolutive Source*. ISBN 3639077970, EBN 978-3639077971. VDM Verlag.

Adler, R. J. (2009). The geometry of random fields: Siam.

Aldroubi, A., & Unser, M. A. (1996). *Wavelets in Medicine and Biology*: CRC.

Alexander, M. E., Baumgartner, R., Windischberger, C., Moser, E., & Somorjai, R. L. (2000). Wavelet domain de-noising of time-courses in MR image sequences. *Magnetic resonance imaging, 18*(9), 1129-1134.

Alkire, M. T., & Haier, R. J. (2001). Correlating in vivo anaesthetic effects with ex vivo receptor density data supports a GABAergic mechanism of action for propofol, but not for isoflurane *British Journal of Anaesthesia, 86*(5), 618-626.

Alkire, M. T., Haier, R. J., & Fallon, J. H. (2000). Toward a unified theory of narcosis: brain imaging evidence for a thalamocortical switch as the neurophysiologic basis of anesthetic-induced unconsciousness. *Consciousness and Cognition, 9*(3), 370-386.

Alkire, M. T., Haier, R. J., Fallon, J. H., & Barker, S. J. (1996). PET imaging of conscious and unconscious verbal memory. *Journal of Consciousness Studies, 3, 5*(6), 448-462.

Alkire, M. T., Haier, R. J., Fallon, J. H., & Cahill, L. (1998). Hippocampal, but not amygdala, activity at encoding correlates with long-term, free recall of nonemotional information. *Proceedings of the National Academy of Sciences, 95*(24), 14506-14510.

Allan, D. W. (1966). Statistics of atomic frequency standards. *Proceedings of the IEEE, 54*(2), 221-230.

Ando, J., Suzuki, A., Yamagata, S., Kijima, N., Maekawa, H., Ono, Y., et al. (2004). Genetic and environmental structure of Cloninger's temperament and character dimensions. *Journal of Personality Disorders, 18*(4), 379-393.

Angelidis, P. A. (1994). MR image compression using a wavelet transform coding algorithm. *Magnetic Resonance Imaging, 12*(7), 1111-1120.

Antoniadis, A. (2007). Wavelet methods in statistics: Some recent developments and their applications. *Statistics Surveys, 1*, 16-55.

Antoniadis, A., & Gijbels, I. (2002). Detecting abrupt changes by wavelet methods. *Journal of Nonparametric Statistics, 14*(1), 7-29.

Antoniadis, A., Leporini, D., & Pesquet, J. C. (2002). Wavelet thresholding for some classes of non-Gaussian noise. *Statistica Neerlandica, 56*(4), 434-453.

Aroian, L. A., & Levene, H. (1950). The effectiveness of quality control charts. *Journal of the American Statistical Association, 45*(252), 520-529.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage, 38*(1), 95-113.

Ashburner, J., & Friston, K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping, 7*(4), 254-266.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry-the methods. *NeuroImage, 11*(6), 805-821.

Attias, H. (1999). Independent factor analysis. *Neural Computation, 11*(5), 803-852.

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 47-78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*(1), 1-22.

Banerjee, A., & Urga, G. (2005). Modelling structural breaks, long memory and stock market volatility: an overview. *Journal of Econometrics, 129*(1-2), 1-34.

Barber, S., Nason, G. P., & Silverman, B. W. (2002). Posterior probability intervals for wavelet thresholding. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 64*(2), 189-205.

Barndorff-Nielsen, O. E., & Cox, D. R. (1989). *Asymptotic techniques for use in statistics*. New York: Chapman and Hall.

Baron, M., & Tartakovsky, A. G. (2006). Asymptotic optimality of change-point detection schemes in general continuous-time models. *Sequential Analysis, 25*(3), 257-296.

Barra, V., & Boire, J. Y. (2000). Tissue segmentation on MR images of the brain by possibilistic clustering on a 3D wavelet representation. *Journal of Magnetic Resonance Imaging, 11*(3), 267-278.

Bartels, A., & Zeki, S. (2004). The chronoarchitecture of the human brain--natural viewing conditions reveal a time-based anatomy of the brain. *NeuroImage, 22*(1), 419-433.

Bartolucci, A., Bae, S., Singh, K., & Griffith, H. R. (2009). An examination of Bayesian statistical approaches to modeling change in cognitive decline in an Alzheimer's disease population. *Mathematics and Computers in Simulation, 80*(3), 561-571.

Bassett, D. S., & Bullmore, E. (2006). Small-world brain networks. *The Neuroscientist, 12*(6), 512-523.

Beckmann, C. F., DeLuca, M., Devlin, J. T., & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1457), 1001-1013.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural computation, 7*(6), 1129-1159.

Bell, A. J., & Sejnowski, T. J. (1996). Learning the higher-order structure of a natural sound. *Network: Computation in Neural Systems, 7*(2), 261-266.

Bell, A. J., & Sejnowski, T. J. (1997). The "independent components"of natural scenes are edge filters. *Vision research, 37*(23), 3327-3338.

Bellman, R., & Dreyfus, S. (1962). *Applied dynamic programming*. New Jersey: Princeton University Press.

Bench, C. J., Frankowiak, R. S. J., & Dolan, R. J. (1995). Changes in reginal cerebral blood flow on recovery from depression. *Psychological Medicine, 25*(2), 247-261.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289-300.

Beran, J. (1994). *Statistics for long-memory processes*: Chapman & Hall/CRC.

Bingham, E., & Hyvärinen, A. (2000). A fast fixed-point algorithm for Independent Component Analysis of complex valued signals. *International Journal of Neural Systems, 10*(1), 1-8.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (2005). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*(4), 537-541.

Bobin, J., Starck, J., Fadili, J., & Moudden, Y. (2007). Sparsity and morphological diversity in blind source separation. *IEEE Transactions on Image Processing, 16*(11), 2662-2674.

Bradshaw, G. A., & Spies, T. A. (1992). Characterizing canopy gap structure in forests using wavelet analysis. *Journal of Ecology, 80*(2), 205-215.

Brammer, M. J. (1998). Multidimensional wavelet analysis of functional magnetic resonance images. *Human Brain Mapping, 6*(5-6), 378-382.

Brattebo, G., Hofoss, D., Flaatten, H., Muri, A. K., Gjerde, S., & Plsek, P. E. (2004). Effect of a scoring system and protocol for sedation on duration of patients' need for ventilator support in a surgical intensive care unit. *British Medical Journal, 13*(3), 203-205.

Bratteli, O., & Jogensen, P. E. T. (2002). Wavelets through a looking glass: The world of the spectrum: Springer.

Breakspear, M., Brammer, M., & Robinson, P. A. (2003). Construction of multivariate surrogate sets from nonlinear data using the wavelet transform. *Physica D: Nonlinear Phenomena, 182*(1-2), 1-22.

Breier, A., Kestler, L., Adler, C., Elman, I., Wiesenfeld, N., Malhotra, A., et al. (1998). Dopamine d2 receptor density and personal detachment in healthy subjects. *Am J Psychiatry, 155*(10), 1440-1442.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth, Belmont.

Brett, M., Penny, W., & Kiebel, S. (2003). *Introduction to random field theory.* Human brain function. London, Academic Press.

Bruce, A., & Gao, H. Y. (1996). *Applied wavelet analysis with S-plus*: Springer New York.

Bruce, A. G., & Gao, H. Y. (1996). Understanding WaveShrink: Variance and bias estimation. *Biometrika, 83*(4), 727-745.

Bullmore, E., Fadili, J., Breakspear, M., Salvador, R., Suckling, J., & Brammer, M. (2003). Wavelets and statistical analysis of functional magnetic resonance images of the human brain. *Statistical methods in medical research, 12*(5), 375-399.

Bullmore, E., Fadili, J., Maxim, V., Sendur, L., Whitcher, B., Suckling, J., et al. (2004). Wavelets and functional magnetic resonance imaging of the human brain. *NeuroImage, 23*, S234-S249.

Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., et al. (2001). Colored noise and computational inference in neurophysiological (fMRI) time series analysis: resampling methods in time and wavelet domains. *Human Brain Mapping, 12*(2), 61-78.

Burrus, C. S., Gopinath, R. A., & Guo, H. (1998). *Introduction to wavelets and wavelet transforms: a primer*: Prentice Hall.

Cai, Z. (2002). Regression quantiles for time series. *Econometric Theory, 18*(01), 169-192.

Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience, 115*(1), 33-42.

Cappelli, C., Penny, R. N., Rea, W., & Reale, M. (2007a). Detecting Multiple Mean Breaks at Unknown Points with Regression Trees. *Mathematics and Computers in Simulation.* 78(2-3), 351-356.

Cappelli, C., Rea, W. S., & Reale, M. (2007b). The Application of Regression Trees to the Detecting of Multiple Structural Breaks in the Mean of a Time Series. *Research Report UCMSD, 4.*

Cardoso, J. F. (1991). Super-symmetric decomposition of the fourth-order cumulant tensor. Blind identification of more sources than sensors. In *Proc. ICASSP'*91, 3109-3112.

Carlstein, E. (1988). Nonparametric change-point estimation. *The Annals of Statistics, 16*(1), 188-197.

Carslaw, H. S. (1930). Introduction to the theory of Fourier's series and integrals and the mathematical theory of the conduction of heat. In *Introduction to the theory of Fourier's series and integrals* (p. Chapter IX): Dover Publications

Čencov, N. N. (1962). Evaluation of an unknown distribution density from observations. *Soviet Mathematics*, 3, 1559-1562.

Chang, S. G., Yu, B., & Vetterli, M. (2000a). Spatially adaptive wavelet thresholding based on context modeling for image denoising. *IEEE Transactions on Image Processing, 9*(9), 1522- 1531.

Chang, S. G., Yu, B., & Vetterli, M. (2000b). Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing, 9*(9), 1532-1546.

Charnbolle, A., De Vore, R. A., Lee, N. Y., & Lucier, B. J. (1998). Nonlinear wavelet image processing: Variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing, 7*(3), 319-335.

Chase, J. G., Rudge, A. D., Shaw, G. M., Wake, G. C., Lee, D., Hudson, I. L., et al. (2004). Modelling and control of the agitation-sedation cycle for critical care patients. *Medical Engineering and Physics, 26*(6), 459-471.

Chen, J., & Gupta, A. K. (1997). Testing and Locating Variance Change points with Application to Stock Prices. *Journal of the American Statistical Association, 92*(438), 739-747.

Chen, J., & Gupta, A. K. (2001). On change point detection and estimation. *Communications in Statistics-Simulation and Computation, 30*(3), 665-697.

Cheng, M. Y., & Raimondo, M. (2008). Kernel methods for optimal change-points estimation in derivatives. *Journal of Computational and Graphical Statistics, 17*(1), 56-75.

Chipman, H. A., Kolaczyk, E. D., & McCulloch, R. E. (1997). Adaptive Bayesian wavelet shrinkage. *Journal of the American Statistical Association, 92*(440).

Chipman, H. A., & Wolfson, L. J. (1999). Prior elicitation in the wavelet domain. *Lecture Notes in Statistics*, 141, 83-94.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the Econometric Society*, 591-605.

Chui, C. K. (1992). *An introduction to wavelets*: Academic Pr.

Cleland, E. E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology & Evolution, 22*(7), 357-365.

Cliff, N. (1987). *Analyzing multivariate data*. San Diego, CA: Harcourt Brace Jovanovich.

Cline, D. B. H., Eubank, R. L., & Speckman, P. L. (1995). Nonparametric estimation of regression curves with discontinuous derivatives. *Journal of Statistical Research, 29*, 17-30 .

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments, 3*, 167-226.

Cloninger, C. R. (1994). Temperament and Personality. *Current Opinion in Neurobiology, 4*, 266-273.

Cloninger, C. R (2002). Functional neuroanatomy and brain imaging of personality and its disorders. In H. A. H. D'haenen, J. A. Den Boer & P. Willner (Eds.), *Textbook of biological Psychiatry* (Vol. 2, pp. 777-790): John Wiley & Sons.

Cloninger, C. R. (1986). A unified biosocial theory of personality and its role in the development of anxiety states. *Psychiatric Developments, 4*(3), 167.

Cloninger, C. R. (1999). The Temperament and Character Inventory-revised. *St Louis, MO: Center for Psychobiology of Personality, Washington University*.

Cloninger, C. R. (2003). Completing the psychobiological architecture of human personality development: Temperament, Character, and Coherence. *Understanding human development: Dialogues with lifespan psychology*, 159-181.

Cloninger, C. R. (2004). *Feeling good: the science of well-being*: Oxford University Press, USA.

Cloninger, C. R. (2005a). Measurement of Personality and its Disorders. In J. M. Oldham, A. E. Skodol & D. S. Bender (Eds.), *The American Psychiatric Publishing textbook of personality disorders* (pp. 143-154): American Psychiatric Publishing.

Cloninger, C. R. (2005b). How does personality influence mortality in the elderly? *Psychosomatic Medicine, 67*(6), 839-840.

Cloninger, C. R. (2006). Differentiating personality deviance, normality, and well-being by the seven-factor psychobiological model. *Differentiating normal and abnormal personality*, 65-81 .

Cloninger, C. R. (2008). The psychobiological theory of temperament and character: Comment on Farmer and Goldberg (2008). *Psychological Assessment, 20*(3), 292-299.

Cloninger, C. R., Przybeck, T. R., & Svrakic, D. M. (1991). The tridimensional personality questionnaire: US normative data. *Psychol Rep, 69*(3 Pt 1), 1047-1057.

Cloninger, C. R., Przybeck, T. R., Svrakic, D. M., & Wetzel, R. D. (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*. St. Louis: Center for Psychobiology of Personality, Washington University.

Cloninger, C. R., & Svrakic, D. M. (1997). Integrative psychobiological approach to psychiatric assessment and treatment. *Psychiatry: Interpersonal and Biological Processes, 60*(2), 120-141.

Cloninger, C. R., Svrakic, D. M., & Przybeck, T. R. (1998). *A psychobio logical model of temperament and character*. Waxmann Verlag.

Clyde, M., & George, E. I. (2000). Flexible empirical Bayes estimation for wavelets. *Journal of the Royal Statistical Society. Series B (Statistical Methodology), 62*(4), 681-698.

Clyde, M., Parmigiani, G., & Vidakovic, B. (1998). Multiple shrinkage and subset selection in wavelets. *Biometrika, 85*(2), 391-402.

Comon, P. (1992). Blind identification in presence of noise.

Costa, P. T., & McCrae, R. R. (1985). The NEO personality inventory manual. (Odessa, FL: Psychological Assessment Resources).

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*: Wiley.

Craigmile, P. F., & Percival, D. B. (2002). Wavelet-based trend detection and estimation. *Encyclopedia of Environmetrics, 4*, 2334-2338.

Crinion, J., Ashburner, J., Leff, A., Brett, M., Price, C., & Friston, K. (2007). Spatial normalization of lesioned brains: performance evaluation and impact on fMRI analyses. *NeuroImage, 37*(3), 866-875.

Crouse, M. S., Nowak, R. D., & Baraniuk, R. G. (1998). Wavelet-based statistical signal processing using Hidden Markov models. *IEEE Transactions on signal processing, 46*(4), 886-902.

Croux, C., & Rousseeuw, P. J. (1992a). A class of high-breakdown scale estimators based on subranges. *Communications in Statistics-Theory and Methods, 21*(7), 1935-1951.

Croux, C., & Rousseeuw, P. J. (1992b). Time-efficient algorithms for two highly robust estimators of scale. *Computational Statistics,* 2, 411-428.

Croux, C., Rousseeuw, P. J., & Hossjer, O. (1994). Generalized S-Estimators. *Journal of the American Statistical Association, 89*(428), 1271-1281.

Cselenyi, Z., Olsson, H., Farde, L., & Guly, B. (2002). Wavelet-aided parametric mapping of cerebral dopamine D2 receptors using the high affinity PET radioligand [11C] FLB 457. *NeuroImage, 17*(1), 47-60.

Csillag, F., & Kabos, S. (2002). Wavelets, boundaries, and the spatial analysis of landscape pattern. *Ecoscience, 9*(2), 177-190.

Curtis, V. A., Dixon, T. A., Morris, R. G., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., et al. (2001). Differential frontal activation in schizophrenia and bipolar illness during verbal fluency. *Journal of Affective Disorders, 66*(2-3), 111-121.

Dale, M. R. T. (2000). *Spatial pattern analysis in plant ecology*: Cambridge Univ Press.

Dale, M. R. T., Dixon, P., Fortin, M. J., Legendre, P., Myers, D. E., & Rosenberg, M. S. (2002). Conceptual and mathematical relationships among methods for spatial analysis. *Ecography*, 558-577.

Daubechies, I. (1992). *Ten lectures on wavelets*: Society for Industrial Mathematics.

Davis, R. A., Lee, T. C. M., & Rodriguez-Yam, G. A. (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association, 101*(473), 223-239.

De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). An introduction to Independent Component Analysis. *Journal of Chemometrics, 14*(3), 123-149.

Delouille, V., Franke, J., & Von Sachs, R. (2001). Nonparametric stochastic regression with design-adapted wavelets. *Sankhy : The Indian Journal of Statistics, Series A, 63*(3), 328-366.

Desco, M., Hernandez, J. A., Santos, A., & Brammer, M. (2001). Multiresolution analysis in fMRI: sensitivity and specificity in the detection of brain activation. *Human Brain Mapping, 14*(1), 16-27.

DeVore, R., & Lucier, B. (1992). Fast wavelet techniques for near-optimal image processing. *Proc. IEEE Military Communications Conference*, New York.

Devous, M. D. (1989). Imaging brain function by single-photon emission computer tomography. *Brain imaging: applications in psychiatry*, 147-234.

Dinov, I. D., Mega, M. S., Thompson, P. M., Woods, R. P., Sumners, D. L., Sowell, E. L., et al. (2002). Quantitative comparison and analysis of brain image registration using frequency-adaptive wavelet shrinkage. *IEEE Transactions on Information Technology in Biomedicine, 6*(1), 73-85.

Doi, H. (2007). Winter flowering phenology of Japanese apricot Prunus mume reflects climate change across Japan. *Climate Research, 34*(2), 99-104.

Donnelly, A., Jones, M. B., & Sweeney, J. (2004). A review of indicators of climate change for use in Ireland. *International Journal of Biometeorology, 49*(1), 1-12.

Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE transactions on information theory, 41*(3), 613-627.

Donoho, D. L., & Johnson, N. L. (1998). Minimax estimation via wavelet shrinkage. *Annals of Statistics, 26*, 879-921.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation via wavelet shrinkage. *Biometrika, 81*, 425-455.

Donoho, D. L., & Johnstone, I. M. (1995a). Adapting to Unknown Smoothness Via Wavelet Shrinkage. *Journal of the american statistical association, 90*(432), 1200-1224.

Donoho, D. L., Johnstone, I. M., & Picard, D. (1995b). Wavelet shrinkage: asymptopia? *Journal of the Royal Statistical Society. Series B (Methodological), 57*(2), 301-369.

Dose, V., & Menzel, A. (2004). Bayesian analysis of climate change impacts in phenology. *Global Change Biology, 10*(2), 259-272.

Dose, V., & Menzel, A. (2006). Bayesian correlation between temperature and blossom onset data. *Global Change Biology, 12*(8), 1451-1459.

Doukhan, P., Oppenheim, G., & Taqqu, M. S. (2003). *Theory and applications of long-range dependence*: Birkhauser.

Dragotti, P. L., & Vetterli, M. (2003). Wavelet footprints: Theory, algorithms, and applications. *IEEE Transactions on Signal Processing, 51*(5), 1306-1323.

Duann, J. R., Hsu, J. L., Wang, H. C., & Jung, T. P. (2003). Identifying Brain Areas with Significant rCBF Differences between Parkinson& Disease and Normal Control Subjects. *Human Brain Mapping Annual Meeting, June*, 8-22, New York City.

Ebmeier, K., Blackwood, D., Murray, C., Souza, V., Walker, M., Dougall, N., et al. (1993). Single-phton emission computed tomography with 99mTc-Exametazime in unmedicated schzophrenic patients. *Biol Psychiatry, 33*, 487-495.

Ebmeier, K. P., Deary, I. J., O'Carroll, R. E., & Prentice, N. (1994). Personality associations with the uptake of the cerebral blood flow marker-super (99m) Tc-Exametazime estimated with Single Photon Emission Tomography. *Personality and Individual Differences, 17*(5), 587-595.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the Bootstrap*, (volume 57 of Monographs on Statistics and Applied Probability). Chapmann & Hall.

Ellinas, J. N., & Manolakis, D. E. (2005, Oct). *Image De-noising based on the Statistical Modeling of Wavelet Coefficients and Quad-Tree Decomposition.* Paper presented at the 5th International Conference on Technology and Automation, Thessaloniki.

Embrechts, P., & Maejima, M. (2002). *Selfsimilar processes*: Princeton Univ Press.

Emir, U. E., Akg, C. B., Ak, A., Ert, A., Sankur, B., & Harmanc, K. (2003). Wavelet denoising vs ICA denoising for functional optical imaging. 1st International IEEE EMBS Conference on Neural Engineering, pp. 384-387,

Engel, J. (1990). Density estimation with Haar series. *Statistics & Probability Letters, 9*(2), 111-117.

Eriksson, C., Omstedt, A., Overland, J. E., Percival, D. B., & Mofjeld, H. O. (2007). Characterizing the European sub-Arctic winter climate since 1500 using ice, temperature, and atmospheric circulation time series. *Journal of Climate, 20*(21), 5316-5334.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993). 3D statistical neuroanatomical models from 305 MRI volumes, In Proc. IEEE—Nuclear Science Symposium and Medical Imaging Conference, 1813-1817.

Fadili, M. J., & Bullmore, E. T. (2002). Wavelet-generalized least squares: a new BLU estimator of linear regression models with 1/f errors. *NeuroImage, 15*(1), 217-232.

Farde, L., Gustavsson, J., & Jonsson, E. (1997). D2 dopamine receptors and personality traits. *Nature, 379* 449–451.

Farmer, A., Mahmood, A., Redman, K., Harris, T., Sadler, S., & McGuffin, P. (2003). A sib-pair study of the Temperament and Character Inventory scales in major depression. *Archives of General Psychiatry, 60*(5), 490-496.

Farmer, R. F., & Goldberg, L. R. (2008a). A psychometric evaluation of the revised Temperament and Character Inventory (TCI-R) and the TCI-40. *Psychological Assessment, 20*(3), 281-291.

Farmer, R. F., & Goldberg, L. R. (2008b). Brain Modules, Personality Layers, Planes of Being, Spiral Structures, and the Equally Implausible Distinction Between TCI-R "Temperament" and "Character" Scales: Reply to Cloninger (2008). *Psychological Assessment, 20*(3), 300-304.

Fergusson, D., Beautrais, A., & Horwood, L. (2003). Vulnerability and resiliency to suicidal behaviours in young people. *Psychological Medicine, 33*, 61-73.

Fisher, W. D. (1958). On grouping for maximum homogeneity. *Journal of the American Statistical Association 53*, 789–798.

Fitch, J., Coyle, E., & Gallagher, N. (1984). Median filtering by threshold decomposition. *IEEE Transactions on Acoustics, Speech and Signal Processing, 32*(6), 1183-1188.

Fitter, A. H., & Fitter, R. S. R. (2002). Rapid changes in flowering time in British plants. *Science, 296*(5573), 1689-1691.

Fitter, A. H., Fitter, R. S. R., Harris, I. T. B., & Williamson, M. H. (1995). Relationships between first flowering date and temperature in the flora of a locality in central England. *Functional Ecology, 9*(1), 55-60.

Flandin, G., & Penny, W. D. (2007). Bayesian fMRI data analysis with sparse spatial basis function priors. *NeuroImage, 34*(3), 1108-1125.

Flynn, J. R. (2007). What is intelligence?: beyond the Flynn effect: Cambridge Univ Press.

Fossati, A., Cloninger, C. R., Villa, D., Borroni, S., Grazioli, F., Giarolli, L., et al. (2007). Reliability and validity of the Italian version of the Temperament and Character Inventory-Revised in an outpatient sample. *Comprehensive Psychiatry, 48*(4), 380-387.

Foufoula-Georgiou, E., & Kumar, P. (1994). *Wavelets in geophysics*: Academic Press.

Frackowiak, R. S. J., Friston, K. J., Dolan, R. J., Frith, C. D., Price, C. J., Ashburner, J. T., et al. (2004). *Human brain function*: Academic Press.

Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., & Mazziotta, J. (1997). *Human Brain function*: Academic Press.

Frackowiak, R. S. J., Friston, K. J., Frith, C. D., Dolan, R. J., Price, C. J., Zeki, S., et al. (2003). *Human Brain Function* (2nd ed.): Academic Press.

Fraser, G. L., & Riker, R. R. (2001a). Advances and controversies in sedating the adult critically ill. *The New York Heath-System Pharmacist, 20*(3), 17-24.

Fraser, G. L., & Riker, R. R. (2001b). Monitoring sedation, agitation, analgesia, and delirium in critically ill adult patients. *Crit Care Clin, 17*(4), 967-987.

Friston, K., Ashburner, J., Poline, J., Frith, C., Heather, J., & Frackowiak, R. (1995a). Spatial registration and normalization of images. *Human Brain Mapping,, 2*, 165-189.

Friston, K., Frith, C., P.F.Liddle, Dolan, R., Lammertsma, A., & Frackowiak, R. (1991). The relationship between global and local changes in PET scans. *Journal of Cerebral Blood Flow and Metabolism, 10*, 458-466.

Friston, K., Holmes, A., Worsley, K., Poline, J., Frith, C., Heather, J., et al. (1995b). Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping 2*, 189-210.

Friston, K., Worsley, K., Poline, J., Frith, C., & Frackowiak, R. (1994). Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping Annual Meeting, 1*, 214-220.

Friston, K. J. (1998). Modes or models: a critique on Independent Component Analysis for fMRI. *Trends in Cognitive Sciences, 2*(10), 373-375.

Friston, K. J., Ashburner, J., & Heather, J. (2003). Statistical Parametric Mapping. *Neuroscience Databases: A Practical Guide*, 215-237.

Friston, K. J., Ashburner, J. T., Kiebel, S. J., Nichols, T. E., & Penny, W. D. (2007). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*: Academic Press.

Friston, K. J., Frith, C. D., Liddle, P. F., Dolan, R. J., Lammertsma, A. A., & Frackowiak, R. S. (1990). The relationship between global and local changes in PET scans. *Journal of cerebral blood flow and metabolism: official journal of the International Society of Cerebral Blood Flow and Metabolism, 10*(4), 458.

Friston, K. J., Frith, C. D., Liddle, P. F., & Frackowiak, R. S. (1991). Comparing functional (PET) images: the assessment of significant change. *Journal of Cerebral Blood flow and Metabolism: 11*(4), 690-699.

Friston, K. J., & Penny, W. (2003). Posterior probability maps and SPMs. *NeuroImage, 19*(3), 1240-1249.

Friston, K. J., Stephan, K. E., Lund, T. E., Morcom, A., & Kiebel, S. (2005). Mixed-effects and fMRI studies. *NeuroImage, 24*(1), 244-252.

Gale, A. (1973). The psychophysiology of individual differences: Studies of extraversion and the EEG. *New Approaches in Psychological Measurement*, 211-256.

Gao, H. Y., & Bruce, A. (1995). WaveShrink with Semisoft Shrinkage. *Statist. Sci. Research Report, 39*.

Gao, W., & Li, B. L. (1993). Wavelet analysis of coherent structures at the atmosphere-forest interface. *Journal of Applied Meteorology, 32*(11), 1717-1725.

Gencay, R., Selcuk, F., & Whitcher, B. (2002). An introduction to wavelets and other filtering methods in finance and economics: Academic Press.

Ghosh, P., & Vaida, F. (2006). Random changepoint modelling of HIV immunologic responses. *Statistics in Medicine, 26*(9), 2074-2087.

Ghosh, S., Beran, J., Heiler, S., Percival, D., & Tinner, W. (2007). Memory, non-stationarity and trend: analysis of environmental time series. *A Changing World*, 223-247.

Gijbels, I., & Goderniaux, A. C. (2004). Bootstrap test for change-points in nonparametric regression. *Journal of Nonparametric Statistics, 16*(3), 591-611.

Gijbels, I., & Goderniaux, A. C. (2005). Data-driven discontinuity detection in derivatives of a regression function. *Communications in Statistics-Theory and Methods, 33*(4), 851-871.

Gijbels, I., Hall, P., & Kneip, A. (1999). On the estimation of jump points in smooth curves. *Annals of the Institute of Statistical Mathematics, 51*(2), 231-251.

Gillespie, N. A., Cloninger, C. R., Heath, A. C., & Martin, N. G. (2003). The genetic and environmental relationship between Cloninger's dimensions of temperament and character. *Personality and Individual Differences, 35*(8), 1931-1946.

Goldman, R. G., Skodol, A. E., McGrath, P. J., & Oldham, J. P. (1994). Relationship between the Tridimensional Personality Questionnaire and DSM-III--R personality traits. *American Journal of Psychiatry, 151*(2), 274-276.

Gordo, O., & Sanz, J. J. (2005). Phenology and climate change: a long-term study in a Mediterranean locality. *Oecologia, 146*(3), 484-495.

Goupillaud, P., Grossmann, A., & Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration, 23*(1), 85-102.

Greeenfield, K. M., Dove, R. A., & Shaw, G. M. (2003). Optimisation of sedation therapy within an intensive care setting, In Proceedings of the World Congress on Medicine In *Physics and Biomedical engineering* Sydney, Australia.

Greenfield, K. M., Dove, R. A., & Shaw, G. M. (2001). *Optimisation of sedation therapy within an intensive care setting.* Paper presented at the Engineering and Physical Sciences in Medicine,30th September, Fremantle, Australia.

Grossman, A., & Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis, 15*, 723-736.

Grucza, R. A., & Goldberg, L. R. (2007). The comparative validity of 11 modern personality inventories: Predictions of behavioral acts, informant reports, and clinical indicators. *Journal of Personality Assessment, 89*(2), 167-187.

Gusnard, D. A., Ollinger, J. M., Shulman, G. L., Cloninger, C. R., Price, J. L., Van Essen, D. C., et al. (2003). Persistence and brain circuitry. *Proceedings of the National Academy of Sciences, 100*(6), 3479-3484.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen, 69*(3), 331-371.

Haier, R. J. (2001). PET studies of learning and individual differences. *Mechanisms of Cognitive Development: Behavioral and Neural Perspectives*, 123-145.

Haier, R. J. (2004). Brain Imaging Studies of Personality: The Slow Revolution. *On the Psychobiology of Personality: Essays in Honor of Marvin Zuckerman*, 331-342.

Haier, R. J., Siegel, B. V., MacLachlan, A., & Soderling, E. (1992). Regional glucose metabolic changes after learning a complex visuospatial/motor task: A Positron Emission Tomographic study. *Brain Research, 570*(1-2), 134-143.

Haier, R. J., Siegel Jr, B. V., Nuechterlein, K. H., Hazlett, E., & Wu, J. C. (1988). Cortical glucose metabolic rate correlates of abstract reasoning and attention studied with Positron Emission Tomography. *Intelligence, 12*, 199-217.

Haier, R. J., White, N. S., & Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence, 31*(5), 429-441.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 383-393.

Hamza, A. B., Luque-Escamilla, P. L., Mart?ez-Aroza, J., & Rom?-Rold, R. (1999). Removing noise and preserving details with relaxed median filters. *Journal of Mathematical Imaging and Vision, 11*(2), 161-177.

Hansen, B. E. (2001). The new econometrics of structural change: dating breaks in the U.S. labor productivity. *Journal of Economic Perspectives, 15*, 117-128.

Hardie, R. C., & Barner, K. E. (1994). Rank conditioned rank selection filters for signal restoration. *Image Processing, 3*, 192–206.

Hartigan, J. A. (1975). *Clustering Algorithms* New York: John Wiley & Sons.

Hayasaka, S., Phan, K. L., Liberzon, I., Worsley, K. J., & Nichols, T. E. (2004). Nonstationary cluster-size inference with random field and permutation methods. *NeuroImage, 22*(2), 676-687.

Hermes, M. (2007). *Personality and Affect / A Psychophysiological Approach* Thesis. Trier University, Tier, Germany.

Hernández, E., & Weiss, G. (1996). A first Course on wavelets. In (pp. 201-224): CRC-Press.

Herrick, D. R. M. (2000). *Wavelet Methods for Curve and Surface Estimation.* Thesis. University of Bristol, Bristol, U.K.

Hettmansperger, T. P., & McKean, J. W. (1998). *Robust Nonparametric Statistical Methods*. London: Arnold.

Hill, I. D. (1976). Algorithm AS100: Normal-Johnson and Johnson-Normal transformationns. *Appl. statist, 25*, 190-192.

Hill, I. D., Hill, R., & Holder, R. L. (1976). Algorithm AS99: Fitting Johnson curves by moments. *Appl. statist, 25*, 180–189.

Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika, 57*(1), 1-17.

Horváth, L., & Kokoszka, P. (1997). The effect of long-range dependence on change-point estimators. *Journal of Statistical Planning and Inference, 64*(1), 57-81.

Hosking, J. R. M., & Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics, 29*(3), 339-349.

Hossein-Zadeh, G. A., Soltanian-Zadeh, H., & Ardekani, B. A. (2003). Multiresolution fMRI activation detection using translation invariant wavelet transform and statistical analysis based on resampling. *IEEE Transactions on Medical Imaging, 22*(3), 302-314.

Hoyer, P. O., & Hyvärinen, A. (2000). ICA features of colour and stereo images. In *Proc. Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA2000)*. pp. 567-572), 19-22 June, Helsinki, Finland.

Hsu, J. L., Duann, J. R., Wang, H. C., & Jung, T. P. (2003). *Assessing rCBF changes in parkinson's disease using independent component analysis*. 4th International Symposium on Independent Component Analysis and Brain Signal Separation. April, Nara, Japan. http://www.wikipedia.org/, W.

Hu, Y. S. (1994). Wavelet approach to change point detection with application to density estimation. Thesis. Texas A& M University, USA.

Huang, C., & Litzenberger, R. H. (1988). *Foundations for financial economics*: North-Holland New York.

Huber, P. (1981). *Robust Statistics*: Wiley.

Huber, P. J. (1964). Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 73-101.

Huber, P. J., & Ronchetti, E. M. (2009). *Robust Statistics*: Wiley-Blackwell.

Hudson, I. L. (2010). Interdisciplinary approaches: towards new statistical methods for phenological studies. *Climatic Change*, 00-28 (In press).

Hudson, I. L. (2010a). Meta-Analysis and Its Application in Phenological Research: a Review and New Statistical Approaches. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research : Methods for Environmental and Climate Change Analysis*, 463-509, Springer Dordrecht.

Hudson, I. L. (2010c).Primary Title: Interdisciplinary approaches: towards new statistical methods for phenological studies, *Climatic Change*, 100(1), 143-171, Url: http://dx.doi.org/10.1007/s10584-010-9859-9, Doi: 10.1007/s10584-010-9859-9

Hudson, I. L. (2010b). Meta-Analysis. In S. S (Ed.), *The Oxford Encyclopedia of Climate and Weather* (pp. 1-14): Oxford University Press, (In Press).

Hudson IL, Barnett A, Keatley MR, Ades PK (2003) Investigation into drivers for flowering in eucalypts: effects of climate on flowering. In: Verbeke G, Moelenberghs G, Aaerts M, Fieuws S (eds) 18th International Workshop on Statistical Modelling, Katholieke Universiteit Leuven, Belgium, pp 195-200

Hudson, I. L., Kang, I., & Keatley, M. R. (2010a). Wavelet Analysis of Flowering and Climatic Niche Identification. In I. L. Hudson & M. R. Keatley (Eds.), *Phenological Research: Methods for Environmental and Climate Change Analysis* (Chapter 17. pp 361-391): Springer Dordrecht.

Hudson, I. L., Kang, I., Rudge, A. D., Chase, J. G., & Shaw, G. M. (2004, Nov 22-23). *Wavelet signatures of agitation and sedation profiles: a preliminary study.* Paper presented at the New Zealand Physics & Engineering in Medicine (NZPEM) Christchurch New Zealand.

Hudson, I. L., & Keatley, M. R. (2010a). *Phenological Research: Methods for Environmental and Climate Change Analysis*: In I.L. Hudson & M.R. Keatley (Eds.). Springer Dordrecht..

Hudson, I. L., & Keatley, M. R. (2010b). Singular Spectrum Analysis: Climatic Niche Identification. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research: Methods for Environmental and Climate Change Analysis*, 393-424, Springer Dordrecht.

Hudson, I. L., Keatley, M. R., & Kang, I. (2010d). Wavelet characterisation of eucalypt flowering and the influence of climate. *Environmental and Ecological Statistics.* (DOI10.1007/s10584-010-98589-9.

Hudson, I. L., Keatley, M. R., & Roberts, A. M. I. (2005). Statistical Methods in Phenological Research. In *International Workshop in Statistical Modelling (IWSM)*. pp. 259-270). Sydney:(Francis, A.R., Matawie, K.M., Oshlack, A. and Smyth G.K.(Eds.)).

Hudson, I. L., Kim, S. W., & Keatley, M. R. (2010b). Climatic influences on the flowering phenology of four eucalypts: a GAMLSS approach. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research: Methods for Environmental and climate Change Analysis*, 209-228, Springer Dordrecht .

Hudson, I. L., Kim, S. W., & Keatley, M. R. (2010c). Modelling the Flowering of Four Eucalypt Species Using New Mixture Transition Distribution Models. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research: Methods for Environmental and Climate Change Analysis*, 299-320 Springer Dordrecht.

Hughes, L. (2003). Climate change and Australia: trends, projections and impacts. *Austral Ecology, 28*(4), 423-443.

Huh, J. (2002). Nonparametric discontinuity point estimation in density or density derivatives. *Journal of the Korean Statistical Society, 31*, 261-276.

Huh, J., & Carriere, K. C. (2002). Estimation of regression functions with a discontinuity in a derivative with local polynomial fits. *Statistics & Probability Letters, 56*(3), 329-343.

Huh, J., & Park, B. U. (2004). Detection of a change point with local polynomial fits for the random design case. *Australian & New Zealand Journal of Statistics, 46*(3), 425-441.

Hurri, J., Hyvärinen, A., Karhunen, J., & Oja, E. (1996). *Image feature extraction using Independent Component Analysis.* Paper presented at theNORSIG'96, Sep, Espoo, Finland.

Hurri, J., Hyvärinen, A., & Oja, E. (1997). Wavelets and natural image statistics. *In Proc. Scandinavian Conf. on Image Analysis '97*. June, Lappenranta, Finland.

Hyafil, L., & Rivest, R. (1976). Constructing optimal binary decision trees is N-P complete. *Information Processing Letters, 5*, 15-17.

Hyvärinen, A. (1998a). New approximations of differential entropy for Independent Component Analysis and Projection Pursuit. *Advances in neural information processing systems*, 273-279.

Hyvärinen, A. (1998b). The FastICA MATLAB toolbox. *Helsinki Univ. of Technology*.

Hyvärinen, A. (1999a). Sparse code shrinkage: Denoising of nongaussian data by maximum likelihood estimation. *Neural Computation, 11*(7), 1739-1768.

Hyvärinen, A. (2001a). Fast ICA by a fixed-point algorithm that maximizes non-Gaussianity. In S. a. E. In Roberts, R (Ed.), *Independent Component Analysis* (pp. 71-94). United Kingdom: Cambridge University Press.

Hyvärinen, A., Cristescu, R., & Oja, E. (1999c). *A fast algorithm for estimating overcomplete ICA bases for image windows.* Paper presented at the In Proc. Int. Joint Conf. on Neural Networks, July, Washington, D.C.

Hyvärinen, A., Karhunen, J., & Oja, E. (2001b). *Independent Component Analysis*: Wiley.

Hyvärinen, A., & Oja, E. (1998a). Independent Component Analysis by general nonlinear Hebbian-like learning rules. *Signal Processing, 64*(3), 301-313.

Hyvärinen, A., & Oja, E. (2000). Independent Component Analysis: algorithms and applications. *Neural Networks, 13*(4-5), 411-430.

Hyvärinen, A., Oja, E., Hoyer, P., & Hurri, J. (1998c). *Image feature extraction by sparse coding and independent component analysis.* Paper presented at the Int. Conf. on Pattern Recognition (ICPR'98), Brisbane, Australia.

Hyvärinen, A., & Pajunen, P. (1999b). Nonlinear Independent Component Analysis: exstence and uniqueness results. *Neural Networks, 12*(3), 429-439.

Hyvärinen, A. (2005), A unifying model for blind separation of independent sources, *Signal Processing*, 85(7), pp. 1419–1427.

Hyvärinen, A., Ramkumar, P., Parkkonen, L., & Hari, R. (2009). Independent Component Analysis of short-time Fourier transforms for spontaneous EEG/MEG analysis. *Neuroimage, 49*(1), 257-271.

Ito, M., Fukuda, M., Suto, T., Uehara, T., & Mikuni, M. (2005). Increased and Decreased Cortical Reactivities in Novelty Seeking and Persistence. A Multichannel Near-Infrared Spectroscopy Study in Healthy Subjects. *Neuropsychobiology, 52*(1), 45-54.

Iyriboz, T. A., Zukoski, M. J., Hopper, K. D., & Stagg, P. L. (1999). A comparison of wavelet and Joint Photographic Experts Group lossy compression methods applied to medical images. *Journal of Digital Imaging, 12*, 14-17.

Jaarsma, A. S., Knoester, H., van Rooyen, F., & Bos, A. P. (2001). Biphasic positive airway pressure ventilation (pev+) in children. *Crit Care Clin, 5*(3), 174-177.

Jaffard, S., Meyer, Y., & Ryan, R. D. (2001). *Wavelets: Tools for science & technology*: Phildelphia, PA, Society for Industrial Mathematics.

Jansen, M. (2001). Noise reduction by wavelet thresholding: Springer USA.

Jansen, M., Malfait, M., & Bultheel, A. (1997). Generalized cross validation for wavelet thresholding. *Signal Processing, 56*(1), 33-44.

Jaruskova, D. (1997). Some problems with application of change-point detection methods to environmental data. *Environmetrics, 8*(5), 469-484.

Jensen, A., & la Cour-Harbo, A. (2001). Ripples in Mathematics : the Discrete Wavelet Transform: Springer.

Johnson, D. L., Wiebe, J. S., Gold, S. M., Andreasen, N. C., Hichwa, R. D., Watkins, G. L., et al. (1999). Cerebral blood flow and personality: A Positron Emission Tomography study. *American Journal of Psychiatry, 156*(2), 252-257.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika, 36*, 149-176.

Johnstone, I. M., & Silverman, B. W. (1997). Wavelet threshold estimators for data with correlated noise. *Journal of the Royal Statistical Society. Series B (Methodological), 59*(2), 319-351.

Jouini, J. (2010). Bootstrap methods for single structural change tests: power versus corrected size and empirical illustration. *Statistical Papers*, 1-25.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., et al. (2000a). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37*(02), 163-178.

Jung, T. P., Makeig, S., Lee, T. W., McKeown, M. J., Brown, G., Bell, A. J., et al. (2000b). *Independent ComponentAanalysis of biomedical signals.* 2nd Int. Workshop on Independent Component Analysis and Signal Separation, 633–644.

Jung, T. P., Makeig, S., McKeonwn, M. J., Lee, T. W., & Sejnowski, T. (2001). *Imaging brain dynamics using Independent Component analysis* Proc IEEE 89(7), 1107-1122.

Kaasinen, V., Maguire, R. P., Kurki, T., Bruck, A., & Rinne, J. O. (2005). Mapping brain structure and personality in late adulthood. *NeuroImage, 24*(2), 315-322.

Kang, I., Hudson, I. L., & Keatley, M. R. (2004). *Wavelet analysis in phenological research-the study of four flowering eucalypt species in relation to climate.* Paper presented at the Intl Biometrics Conference (IBC 04), July, Cairns, Australia.

Kang, I., Hudson, I. L., Rudge, A. D., & Chase, J. G. (2005). *Wavelet signature of Agitation-Sedation profiles of ICU patients.* Francis, A.R., Matawie, K.M., Oshlack, A. and Smyth G.K.(Eds.), International Workshop in Statistical Modelling *IWSM*, July, Sydney, 293-296.

Kapetanios, G., Tzavalis, E., & Mary, Q. (2004). Modelling Structural Breaks: Feburary, Queen Mary University.

Kaplan, A. Y., & Shishkin, S. L. (2000). Application of the change-point analysis to the investigation of the brain electrical activity In B. E. Brodsky & B. S. Darkhovsky (Eds.), *Non-parametric Statistical Diagnosis. Problems and Methods. Dordrecht:* Kluwer Academic Publishers , Chapter 7, 338-388..

Katul, G., Lai, C. T., Sch äer, K., Vidakovic, B., Albertson, J., Ellsworth, D., et al. (2001). Multiscale analysis of vegetation surface fluxes: from seconds to years. *Advances in Water Resources, 24*(9-10), 1119-1132.

Kaur, L., Gupta, S., & Chauhan, R. C. (2002). *Image denoising using wavelet thresholding.*

Keatley, M. R., Fletcher, T. D., Hudson, I. L., & Ades, P. K. (2004). *Shifts in the date of flowering commencement in some Australian Plants.* Paper presented at the Biometeorology and AeroBiology, Vancouver, Canada.

Keatley, M. R., & Hudson, I. L. (2007). *Shifts inflowering dates of Aust related to climate..* In: Oxley L, Kulasiri D (eds) MODSIM 2007 International Congress on Modelling and Simulation Land, Water and Environmental Management: Integrated Systems for Sustainability Modelling and Simulation Society of Australia and New Zealand Christchurch, New Zealand, 504-510.

Keatley, M. R., & Hudson, I. L. (2010). Introduction and Overview. In I. L. Hudson & M. R. Keatley (Eds.), *Phenological Research: Methods for Environmental and Climate Change Analysis* ,= 1-22, Springer, Dordrecht.

Keatley, M. R., Hudson, I. L., & Fletcher, T. D. (2004). Long-term flowering synchrony of box-ironbark eucalypts. *Australian Journal of Botany, 52*(1), 47-54.

Khattree, R., & Naik, D. N. (2008). *Computational Methods in Biomedical Research*: Chapman & Hall.

Kiebel, S., & Friston, K. J. (2002). Anatomically informed basis functions in multisubject studies. *Human Brain Mapping, 16*(1), 36-46.

Kiebel, S., & Holmes, A. P. (2003). The general linear model. In *Human brain function* (2nd edition ed., pp. 725-760): Academic Press.

Kiebel, S. J., Poline, J. B., Friston, K. J., Holmes, A. P., & Worsley, K. J. (1999). Robust smoothness estimation in statistical parametric maps using standardized residuals from the general linear model. *NeuroImage, 10*(6), 756-766.

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M. C., et al. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage, 46*(3), 786-802.

Koch, E. (2010). Global Framework for Data CollectionData Bases, Data Availability, Future Networks, Online Databases. In I. L. Hudson & M. R. Keatley (Eds.), *Phenological Research Methods for Environmental and Climate Change Analysis* 23-61, Springer Dordrecht.

Koch, M., & Marković, D. (2007). Evidences for climate change in Germany over the 20th century from the stochastic analysis of hydro-meteorological time-series. In L. Oxley & D. Kulasiri (Eds.). In *MODSIM 2007 International Congress on Modelling and Simulation Land, Water and Environmental Management: Integrated Systems for Sustainability Modelling and Simulation Society of Australia and New Zealand.* . 596-602. Christchurch, New Zealand.

Kolb, B., and Whishaw, Ian Q. (Cartographer). (1990). *Fundamentals of Human Neuropsychology*, 4th edn. New York: W.H. Freeman and Co.

Koop, G., & Potter, S. (2004). Forecasting and estimating multiple change-point models with an unknown number of change-points. *Federal Reserve Bank of New York Staff Report, 196.*

Korostel, A. P., & Tsybakov, A. B. (1993). Minimax theory of image reconstruction, volume 82 of Lecture Notes in Statistics. (Springer-Verlag, New York. ISBN 0-387-94028-6).

Kotz, S., Johnson, N. L., & Read, C. B. (1982). *Encyclopedia of Statistical Sciences* (Vol. 3). New York: John Wiley & Sons.

Krause, D. N., Duckles, S. P., & Pelligrino, D. A. (2006). Influence of sex steroid hormones on cerebrovascular function. *Journal ofAapplied Physiology, 101*(4), 1252-1261.

Kress, J. P., Pohlman, A. S., & Hall, J. B. (2002). Sedation and analgesia in the intensive care unit. *Am J Respir Crit Care Med, 166*(8), 1024-1028.

Kriegeskorte, N., & Bandettini, P. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. *NeuroImage, 38*(4), 649-662.

Kumar, P. Role of coherent structures in the stochastic-dynamic variability of precipitation. *Journal of Geophysical Research-Atmospheres, 101*(D21), 26, 393-404.

Kumar, P., & Foufoula-Georgiou, E. (1993). A multicomponent decomposition of spatial rainfall fields. 1. Segregation of large- and small-scale features using wavelet transforms. *Water Resources Research, 29*(8), 2515-2532.

Kybic, J., Thevenaz, P., Nirkko, A., & Unser, M. (2000). Unwarping of unidirectionally distorted EPI images. *IEEE Transactions on Medical Imaging, 19*(2), 80-93.

Laine, A. F. (2000). Wavelets in Temporal and Spatial Processing of Biomedical Images. *Annual Review of Biomedical Engineering, 2*(1), 511-550.

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach Atlas labels for functional brain mapping Web address for this project: http://nc. uthscsa. edu/projects/tnlairachdaemon. html. *Human Brain Mapping, 10*(3), 120-131.

Lark, R. M., & Webster, R. (2008). Analysis and elucidation of soil variation using wavelets. *European Journal of Soil Science, 50*(2), 185-206.

Lathauwer, L. D., Moor, B. De, and J. Vandewalle. (1996). *A technique for higher-order-only blind source separation.* Paper presented at the ICONIP, Hong Kong.

Lee, D. S., Rudge, A. D., Chase, J. G., & Shaw, G. M. (2005). A new model validation tool using kernel regression and density estimation. *Computer Methods and Programs in Biomedicine, 80*, 75-87.

Lee, D. S., Rudge, A. D., Chase, J. G., Shaw, G. M., & Johnston, L. (2003). Dynamic model assessment using a probability band for local linear kernel regression, with an application in agitation–sedation modeling. Paper presented at the New Zealand Statistics Association (NZSA), Massey University, New Zealnd.

Lee, T. W. (1998). *Independent component analysis: theory and applications*: Kluwer academic publishers Boston, USA.

Lee, T. W., & Lewicki, M. S. (2002). Unsupervised image classification, segmentation, and enhancement using ICA mixture models. *IEEE Transactions on Image Processing, 11*(3), 270-279.

Lewicki, M., & Olshausen, B. (1998a). *Inferring sparse, overcomplete image codes using an efficient coding framework.* Paper presented at the In Advances in Neural Information Processing System (Proc. NIPS*97), 10, 815-821, MIT Press. .

Lewicki, M., & Sejnowski, T. J. (1998b). *Learing overcomplete representations.* Paper presented at the In Advances in Neural Information Processing 10 (Proc. NIPS*97), 556-562, MIT Press.

Lio, P., & Vannucci, M. (2000). Wavelet change-point prediction of transmembrane proteins. *Bioinformatics, 16*(4), 376-382.

Lord, G. J., Pardo-Iguzquiza, E., & Smith, I. M. (2000). *A practical guide to wavelets for metrology*: National Physical Laboratory.

Lu, X., Liu, R., Liu, J., & Liang, S. (2007). Removal of noise by wavelet methods to generate high quality temporal data of terrestrial MODIS products. *Photogrammetric Engineering and Remote Sensing, 73*(10), 1129-1140.

MacGillivray, F., Hudson, I. L., & Lowe, A. J. (2010). Herbarium Collections and Photographic Images: Alternative Data Sources for Phenological Research. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research: Methods for Environmental and Climate Change Analysis*, 425-461, Springer Dordrecht.

Maguire, B. A., Pearson, E. S., & Wynn, A. H. A. (1952). The Time Intervals Between Industrial Accidents. *Biometrika, 38*, 168-180.

Mahmood, A. K., & Selin, A. (2006). *Spatially adaptive wavelet thresholding for image watermarking*, IEEE International Conference on Multimedia & Expo (ICME), July, Toronto, Canada.

Makeig, S., Westerfield, M., Jung, T. P., Covington, J., Townsend, J., Sejnowski, T. J., et al. (1999). Functionally independent components of the late positive event-related potential during visual spatial attention. *Journal of Neuroscience, 19*(7), 2665-2680.

Maldjian, J. A., Liu, W. C., Hirschorn, D., Murthy, R., & Semanczuk, W. (1997). Wavelet transform-based image compression for transmission of MR data. *American Journal of Roentgenology, 169*, 23-26.

Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern analysis and Machine Intelligence, 11*, 674-693.

Mallat, S., & Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Transactions on Tnformation Theory, 38*(2 Part 2), 617-643.

Mallat, S. G. (1998). *A wavelet tour of signal processing*: Academic Press.

Mallat, S. G. (2009). A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way: Academic Press.

Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM review*, 422-437.

Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust Statistics: Theory and Methods*: Wiley Series in Probability and Statistics.

Mathew, R. J., Weinman, M. L., & Barr, D. L. (1984). Personality and regional cerebral blood flow. *British Journal of Psychiatry, 144*, 529-532.

Mathworks lnc., M. *Image Processing Toolbox*, from http://www.mathworks.com/products/wavelet (accessed 3/5/2010).

Matsuda, H. (2001). Cerebral blood flow and metabolic abnormalities in Alzheimer's diseases. *Ann Nucl Med, 15*(2), 85-92.

Matthews, F., Pearlmutter, B. A., Ward, T. E., Soraghan, C., & Markham, C. (2008). Hemodynamics for Brain-Computer lnterfaces. *IEEE Signal Processing Magazine, 25*(1), 87-94.

Mayberg, H., Liotti, M., Brannan, S., McGinnis, S., Mahurin, R., Jerabek, P., et al. (1999). Reciprocal Limbic-cortical function and negative mood: converging PET findings in depression and normal sadness. *Am J Psychiatr, 156*(5), 675-682.

McCoy, E. J., Percival, D. B., & Walden, A. T. (1995). *Spectrum estimation via wavelet thresholding of multitaper estimators* lmperial College, London.

McCoy, E. J., & Walden, A. T. (1996). Wavelet analysis and synthesis of stationary long-memory processes. *Journal of Computational and Graphical Statistics, 5*(1), 26-56.

McKeown, M. J., Jung, T.-P., Makeig, S., Brown, G. G., Lee, T.-W., Kindermann, S. S., et al. (1998a). *Spatially independent activity patterns in functional MRI data during the stroop color-naming task*. Paper presented at the Proc. Natl. Acad. Sci, USA.

McKeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., & Sejnowski, T. J. (1998b). Analysis of fMRI data by blind separation into independent spatial components. *Hum. Brain Map, 6*, 160-188.

McKeown, M. J., & Sejnowski, T. J. (1998c). Independent Component Analysis of fMRI data: Examining the assumptions. *Hum. Brain Map, 6*, 368-372.

Meyer, F. G. (2003). Wavelet-based estimation of a semiparametric generalized linear model of fMRI time-series. *IEEE Transactions on Medical Imaging, 22*(3), 315-322.

Meyer, J., Wilson, A., Ginovat, N., Goulding, V., Hussey, D., Hood, K., et al. (2001). Occupancy of serotonin transporters by paroxetine and citalopram during tratment of depression a [(11)C]DASB PET imaging study. *Am J Psychiatry, 158*(11), 1843-1849.

Mihcak, M. K., Kozintsev, I., & Ramchandran, K. (1999). Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising. Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing, vol. 6, Mar. 3253–3256.

Mihcak, M. K., Kozintsev, I., & Ramchandran, K. (1999). Spatially adaptive statistical modelling of wavelet image coefficients and its application to denoising. In *Acoustics, Speech and Signal Processing*. Chicago, IL, USA.

Minto, C. F., Schnider, T. W., Short, T. G., Gregg, K. M., Gentilini, A., & Shafer, S. L. (2000). Response surface model for anesthetic drug interactions. *Anesthesiology, 92*(6), 1603-1616.

Misiti, M., Misiti, Y., Oppenheim, G., & Poggi, J. M. (1997). *Wavelet toolbox: for use with MATLAB: User's Guide: Version 3*: Math Works.

Mohanty, S. D., & Jimenez, A. (2005). Non-parametric change point detection. *Classical and Quantum Gravity, 22*, S1233-S1241.

Mohideen, S. K., Perumal, S. A., & Sathik, M. M. (2008). Image de-noising using discrete wavelet transform. *IJCSNS, 8*(1), 213-216.

Molenaar, P. C. M., Huizenga, H. M., & Nesselroade, J. R. (2003). The Relationship Between the Structure of Interindividual and Intraindividual Variability: A Theoretical and Empirical Vindication of Developmental Systems Theory. *Understanding human development: Dialogues with lifespan psychology*, 339-360.

Møller, A. P., Rubolini, D., & Lehikoinen, E. (2008). Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences, 105*(42), 16195-16200.

Mondal, D., & Percival, D. B. (2009). M-Estimation of Wavelet Variance, Ann. Inst. Statist. Math.(In Press).

Morabito, A. N., Percival, D. B., Sahr, J. D., Berkowitz, Z. M. P., & Vertatschitsch, L. E. (2008). Ricean parameter estimation using phase information in low SNR environments. *IEEE Communications Letters, 12*(4), 244-246.

Morlet, J. (1983). Sampling theory and wave propagation. *Issues in Acoustic Signal/Image Processing and Recognition, 1*, 233-261.

Motwani, M. C., Gadiya, M. C., & Motwani, R. C. (2004). *Survey of Image Denoising Techniques.* Paper presented at the GSP Santa Clara Convention Centre, September, Snata Clara, CA, 27-30.

Moulines, E., Cardoso, J. F., & Gassiat, E. (1997). *Maximum likelihood for blind separation and deconvolution of noisy signals using mixture models.* In Proc. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 3617–3620.

Müeller, H. G. (1992). Change-points in nonparametric regression analysis. *The Annals of Statistics, 20*(2), 737-761.

Müeller, H. G. (1998). Non-parametric models for non-smooth functions. In B. Szyszkowicz (Ed.), *Asymptotic methods in probability and statistics: A Volume in Honour of Miklós Csörgő* (pp. 595–609): North-Holland.

Mulder, R. T., & Joyce, P. R. (1997). Temperament and the structure of personality disorder symptoms. *Psychological Medicine, 27*(01), 99-106.

Müler, K., Lohmann, G., Zysset, S., & Von Cramon, D. Y. (2002). Wavelet statistics of functional MRI data and the general linear model. *Journal of Magnetic Resonance Imaging, 17*(1), 20-30.

Nason, G., Kovac, A., & Machler, M. (2004). Wavethresh: Software to perform wavelet statistics and transforms. *URL http://CRAN. R-project. org/src/contrib/Descriptions/wavethresh. html.*

Nason, G. P. (1993). The Wavethresh package; wavelet transform and thresholding software for S. *Available from the StatLib archive.*

Nason, G. P. (1995). Choice of the threshold parameter in wavelet function estimation. *Wavelets and statistics, 103*, 261-280.

Nason, G. P. (2008). *Wavelet methods in statistics with R*: Springer Verlag.

Nason, G. P., & Silverman, B. W. (1994). The discrete wavelet transform in S. *Journal of Computational and Graphical Statistics, 3*(2), 163-191.

Neumann, M. H. (1997). Optimal change-point estimation in inverse problems. *Scandinavian Journal of Statistics, 24*(4), 503-521.

Nichols-Page, G. A., Percival, D. B., Reinhalla, P. G., & Riley, J. J. (2008). Should structure functions be used to estimate power laws in turbulence? A comparative study. *Physica D, 237*, 665-677.

Nolte, J. M., St Louis. (1981). *The Human Brain and Introduction to its Functional Anatomy*: Third ed. Elsevier Health Sciences.

Nowak, D. J., Eisner, R. L., & Fajman, W. A. (1986). Distance-weighted backprojection: a SPECT reconstruction technique. *Radiology, 159*(2), 531-536.

O'Gorman, R. L., Kumari, V., Williams, S. C. R., Zelaya, F. O., Connor, S. E. J., Alsop, D. C., et al. (2006). Personality factors correlate with regional cerebral perfusion. *NeuroImage, 31*(2), 489-495.

Ogden, R. T. (1997). Essential Wavelets for Statistical Applications and Data Analysis. Boston: Birkhäuser.

Ogden, R. T. (1999). *Wavelets in the Bayesian change-point analysis*. Technical Report, Department of Statistics, University of South Carolina.

Ogden, T., & Parzen, E. (1996a). Change-point approach to data analytic wavelet thresholding. *Statistics and Computing, 6*(2), 93-99.

Ogden, T., & Parzen, E. (1996b). Data dependent wavelet thresholding in nonparametirc regression with change-point applications. *Computational Statistics and Data Analysis, 22*, 53-70.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*, 607-609.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1. *Vision Research, 37*, 3311-3325.

Oppenheim, A. V., & Schafer, R. W. (1989). *Discrete-Time signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall.

Overland, J. E., Spillane, M. C., Percival, D. B., Wang, M., & Mofjeld, H. O. (2004). Seasonal and regional variation of pan-Arctic surface air temperature over the instrumental record. *Journal of Climate, 17*(17), 3263-3282.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika, 41*(1/2), 100-115.

Page, E. S. (1955a). A test for a change in a parameter occurring at an unknown point. *Biometrika, 42*(3), 523-527.

Page, E. S. (1955b). A test for a change in a parameter occurring at an unknown point. Vol. 42, pp. 523-527) (Biometrika Trust).

Palma, W. (2007). Long-memory Time Series: Theory and Methods: Wiley-Interscience.

Paluš, M., Novotná, D., & Tichavský, P. (2005). Shifts of seasons at the European mid-latitudes: Natural fluctuations correlated with the North Atlantic Oscillation. *Geophys. Res. Lett, 32*, L12805.

Papoulis, A. (1991). Probability, Random Variables, and Stochastic Processes, Chapter 15. (McGraw-Hill).

Papoulis, A., & Pillai, S. U. (2002). *Probability, random variables, and stochastic processes* (4th ed.). Boston: McGraw-Hill.

Paris, J. (2005). Neurobiological dimensional models of personality: A review of the models of Cloninger, Depue, and Siever. *Journal of Personality Disorders, 19*(2), 156-170.

Park, C., & Kim, W. C. (2006). Wavelet estimation of a regression function with a sharp change point in a random design. *Journal of Statistical Planning and Inference, 136*(7), 2381-2394.

Park, C. W., & Kim, W. C. (2004). Estimation of a regression function with a sharp change point using boundary wavelets. *Statistics and Probability Letters, 66*(4), 435-448.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Reviews of Ecology, Evolution and Systematics*, 37, 637–669.

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature, 421*(6918), 37-42.

Penny, W., Flandin, G., & Trujillo-Barreto, N. (2007). Bayesian comparison of spatially regularised general linear models. *Human Brain Mapping, 28*(4), 275-293.

Penny, W., & Friston, K. (2003). Mixtures of general linear models for functional neuroimaging. *IEEE Transactions on Medical Imaging, 22*(4), 504-514.

Penny, W. D., Kilner, J., & Blankenburg, F. (2007). Robust Bayesian general Linear Models. *NeuroImage, 36*(3), 661-671.

Percival, D. B. (1995). On estimation of the wavelet variance. *Biometrika, 82*(3), 619-631.

Percival, D. B., & Guttorp, P. (1994). Long-memory processes, the Allan variance and wavelets. *Wavelet Analysis and Its Applications*, 325–344.

Percival, D. B., & Mofjeld, H. O. (1997). Analysis of subtidal coastal sea level fluctuations using wavelets. *Journal of the American Statistical Association, 92*(439), 868-880.

Percival, D. B., Overland, J. E., & Mofjeld, H. O. (2004). Modeling North Pacific climate time series. *Time Series Analysis and Applications to Geophysical Systems*, 151-167.

Percival, D. B., & Rothrock, D. A. (2005). "Eyeballing"Trends in Climate Time Series: A Cautionary Note. *J. Climate, 18*(6), 886-891.

Percival, D. B., Rothrock, D. A., Thorndike, A. S., & Gneiting, T. (2008). The variance of mean sea-ice thickness: Effect of long-range dependence. *J. Geophys. Res, 113*, 1958-1970.

Percival, D. B., Sardy, S., & Davison, A. C. (2001). Wavestrapping time series: Adaptive wavelet-based bootstrapping. *Nonlinear and Nonstationary Signal Processing*, 442-470.

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge, England: Cambridge University Press.

Perron, P. (2006). Dealing with structural breaks. *Palgrave Handbook of Econometrics, 1*, 278-252.

Pickands III, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics, 3*(1), 119-131.

Poline, J. B., Kherif, F., & Penny, W. (2004). Contrasts and classical inference. *Human Brain Function, second edR Academic Press, London*, 761-778.

Poline, J. B., Worsley, K. J., Holmes, A. P., Frackowiak, R. S. J., & Friston, K. J. (1995). Estimating smoothness in statistical parametric maps: variability of p values. *Journal of computer assisted tomography, 19*(5), 788-796.

Powell, G. E. (1981). A survey of the effects of brain lesions upon personality. *HJ Eysenck (Hg.), A Model for Personality*, 65-87.

Prohovnik, I. (1993). SPECT imaging of cerebral physiology. In Oldham, Riba, M., ans Tasman, A. (Ed.), *Review of Psychiatry* (Vol. 12, pp. 421-460). Washington DC: American Psychiatric Press Inc.

Raimondo, M. (1998). Minimax estimation of sharp change points. *The Annals of Statistics, 26*(4), 1379-1397.

Raimondo, M. (2002). Wavelet shrinkage via peaks over threshold. *Statistics on the Internet (InterStat)*, 1-19.

Raimondo, M., & Tajvidi, N. (2004). A peaks over threshold model for change-point detection by wavelets. *Statistica Sinica, 14*(2), 395-412.

Ramanaiah, N. V., Rielage, J. K., & Cheng, Y. (2002). Cloninger's temperament and character inventory and the NEO Five-Factor Inventory. *Psychological reports, 90*(3 Pt 2), 1059-1063.

Rammig, A., Jonas, T., Zimmermann, N. E., & Rixen, C. (2009). Changes in alpine plant growth under future climate conditions. *Changes, 6*, 10817-10847.

Rammig, A., Jösson, A. M., Hickler, T., Smith, B., B äring, L., & Sykes, M. T. (2010). Impacts of changing frost regimes on Swedish forests: Incorporating cold hardiness in a regional ecosystem model. *Ecological Modelling*. 221 (2), 303-313.

Ramsay, M. A., Savege, T. M., Simpson, B. R., & Goodwin, R. (1974). Controlled sedation with alphaxalone-alphadolone. *Br Med J, 2*(920), 656-659.

Raz, J., Dickerson, L., & Turetsky, B. (1999). A wavelet packet model of evoked potentials. *Ann Arbor, 1001*, 48109-42029.

Reiss, M. (2005). Estimating the delay time in affine stochastic delay differential equations. International Journal of Wavelets, Multiresolution and Information Processing (IJWMIP), *2*, 525-544.

Riker, R. R., Picard, J. T., & Fraser, G. L. (1999). Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. *Critical care medicine, 27*(7), 1325-1329.

Ripley, B. (2005). *tree: Classification and Regression Trees*. R package version, 1.0-19.

Roberts, A. M. I. (2008). Exploring relationships between phenological and weather data using smoothing. *International Journal of Biometeorology, 52*(6), 463-470.

Roberts, A. M. I. (2010). Smoothing Methods. In I. L. Hudson & M. R. Keatley (Eds.), *Phenological Research: Methods for Environmental and Climate Change Analysis* (pp. 255-269): Springer Dordrecht.

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2001). The kids are alright: Growth and stability in personality development from adolescence to adulthood. *Journal of Personality and Social Psychology, 81*(4), 670-683.

Roberts, S., & Everson, R. (2001). *Independent Component Analysis*. United Kingdom: Cambridge University Press.

Root, T. L., Price, J. T., Hall, K. R., Schneider, S. H., Rosenzweig, C., & Pounds, J. A. (2003). Fingerprints of global warming on wild animals and plants. *Nature, 421*(6918), 57-60.

Rorabaugh, C. B. (1998). *DSP primer*: McGraw-Hill Professional.

Rosenzweig, C., Karoly, D., Vicarelli, M., Neofotis, P., Wu, Q., Casassa, G., et al. (2008). Attributing physical and biological impacts to anthropogenic climate change. *Nature, 453*(7193), 353-357.

Rothrock, D. A., Percival, D. B., & Wensnahan, M. (2008). The decline in arctic sea-ice thickness: Separating the spatial, annual, and interannual variability in a quarter century of submarine data. *Journal of Geophysical Research, 113*, C05003.

Rousseeuw, P. J., & Croux, C. (1992). Explicit scale estimators with high breakdown point. *L*, L_1-Statistical Analysis and Related Methods ,77-92.

Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 1273-1283.

Rudge, A. D., Chase, J. G., Shaw, G. M., & Lee, D. (2006a). Physiological modelling of agitation-sedation dynamics including endogenous agitation reduction. *Medical engineering & physics, 28*(7), 629-638.

Rudge, A. D., Chase, J. G., Shaw, G. M., & Lee, D. (2006b). Physiological modelling of agitation-sedation dynamics. *Medical Engineering and Physics, 28*(1), 49-59.

Rudge, A. D., Chase, J. G., Shaw, G. M., Lee, D., Wake, G. C., Hudson, I. L., et al. (2005). Impact of control on agitation-sedation dynamics. *Control Engineering Practice, 13*(9), 1139-1149.

Rudge, A. D., Chase, J. G., Shaw, G. M., & Wake, G. C. (2003). Improved agitation management in critically ill patients via feedback control of sedation administration. *In Proc World Congress on Medical Physics and Biomed Eng*. August, Sydney, Australia.

Ruggeri, F., & Vidakovic, B. (1998). A Bayesian decision theoretic approach to wavelet thresholding. *J. Amer. Statist. Assoc, 93*(441), 173-179.

Ruttimann, U. E., Unser, M., Rawlings, R. R., Rio, D., Ramsey, N. F., Mattay, V. S., et al. (1998). Statistical analysis of functional MRI data in the wavelet domain. *IEEE Transactions on Medical Imaging, 17*(2), 142-154.

Säckinger, E. (2005). Broadband circuits for optical fiber communication: Wiley-Interscience.

Saito, N. (1994). Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion. *Wavelets in Geophysics*, 299324.

Salmond, C. H., Ashburner, J., Vargha-Khadem, F., Connelly, A., Gadian, D. G., & Friston, K. J. (2002). The precision of anatomical normalization in the medial temporal lobe using spatial basis functions. *NeuroImage, 17*(1), 507-512.

Salvador, R., Suckling, J., Schwarzbauer, C., & Bullmore, E. (2005). Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1457), 937-946.

Schroeder, D. J. (2000). *Astronomical optics*: Academic Press.

Sen, A., & Srivastava, M. S. (1975). Some one-sided tests for change in level. *Technometrics, 17*(1), 61-64.

Serroukh, A., & Walden, A. T. (2001a). Wavelet scale analysis of bivariate time series I: motivation and estimation. *Journal of Nonparametric Statistics, 13*(1), 1-36.

Serroukh, A., & Walden, A. T. (2001b). Wavelet Scale Analysis of Bivariate, Time Series II: Statistical Properties for Linear Processes. *Journal of Nonparametric Statistics, 13*(1), 37-56.

Sessler, C. N., Gosnell, M. S., Grap, M. J., Brophy, G. M., O'neal, P. V., Keane, K. A., et al. (2002). The Richmond agitation-sedation scale: validity and reliability in adult intensive care unit patients. *Am J Respir Crit Care Med, 166*(10), 1338-1344.

Sharifzadeh, M., Azmoodeh, F., & Shahabi, C. (2005). Change detection in time series data using wavelet footprints. *Lecture Notes in Computer Science, 3633*, 127-144.

Shaw, G. M., Chase, J. G., Wong, J., Lin, J., Lotz, T., Le Compte, A. J., et al. (2006). Rethinking glycaemic control in critical illness-from concept to clinical practice change. *Critical Care and Resuscitation, 8*(2), 90-99.

Shaw, G. M., Dove, R. A., Greenfield, K. M., Rudge, A. D., & Chase, J. G. (2003). A computerized approach to sedation administration in critically ill patients. In *ANZICS/ACCCN 28th Australia and New Zealand ASM on Intensive Care*, April, Dunedin, New Zealand.

Shimizu, S. a. K., Y. (2003). *Examination of independence in Independent Component Analysis*. Paper presented at the New Developments in Psychometrics (Proc. IMPS2001), July, Osaka, Tokyo.

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis (pp. 45-47). New York: Chapman & Hall.

Silverman, B. W., & Vassilicos, J. C. (2000). *Wavelets: the key to intermittent information?* : Oxford University Press, USA.

Simoncelli, E. P., & Adelson, E. H. (1996). Noise removal via Bayesian wavelet coring. Proc. IEEE Int. Conf. Image Processing, vol. 1, pp. 379–382.

Sparks, T., & Tryjanowski, P. (2010). Regression and Causality. In I.L. Hudson & M.R. Keatley (Eds.). *Phenological Research: : Methods for Environmental and Climate Change Analysis*, 123-145, Springer Dordrecht.

Sparks, T. H., Jeffree, E. P., & Jeffree, C. E. (2000). An examination of the relationship between flowering times and temperature at the national scale using long-term phenological records from the UK. *International Journal of Biometeorology, 44*(2), 82-87.

Sparks, T. H., & Tryjanowski, P. (2005). The detection of climate impacts: some methodological considerations. *International Journal of Climatology, 25*(2), 271-277.

Stein, C. M. (1981). Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, 1135-1151.

Stelmack, R. M. (2004). On the Psychobiology of Personality: Essays in Honor of Marvin Zuckerman: Elsevier, Amsterdam; Tokyo.

Stenberg, G., Risberg, J., Warkentin, S., & Rosen, I. (1990). Regional patterns of cortical blood flow distinguish extraverts from introverts. *Personality and Individual Differences, 11*(7), 663-673.

Stenberg, G., Wendt, P. E., & Risberg, J. (1993). Regional cerebral blood flow and extraversion. *Personality and individual differences, 15*, 547-547.

Stenseth, N. C., Mysterud, A., Ottersen, G., Hurrell, J. W., Chan, K. S., & Lima, M. (2002). Ecological effects of climate fluctuations. *Science, 297*(5585), 1292-1296.

Stökli, R., Rutishauser, T., Dragoni, D., O'eefe, J., Thornton, P. E., Jolly, M., et al. (2008). Remote sensing data assimilation for a prognostic phenology model. *J. Geophys. Res,* 113-131.

Stone, J. V. (2004). Independent Component Analysis: A Tutorial Introduction: MIT Press.

Stone, J.V. (2005). *A Brief Introduction to Independent Componen Analysis* in Encyclopedia of Statistics in Behavioral Science, Volume 2, pp. 907–912, Editors Brian S. Everitt & David C. Howell, John Wiley & Sons, Ltd, Chichester, 2005 ISBN 978-0-470-86080-9.

Strack, S. (2006). Differentiating normal and abnormal personality: Springer Publishing Co.

Strelau, J. (1983). *Temperament Personality Activity*: Academic Press.

Stuart, A., & Ord, J. K. (1994). Kendall's Advanced Theory of Statistics,: Distribution Theory. (Vol. 1), London: Edward Arnold.

Su, X. G., Wang, M., & Fan, J. J. (2004). Maximum likelihood regression trees. *Journal of Computational and Graphical Statistics, 13*, 586-598.

Suckling, J., & Bullmore, E. (2004). Permutation tests for factorially designed neuroimaging experiments. *Human Brain Mapping, 22*(3), 193-205.

Sugiura, M., Kawashima, R., Nakagawa, M., Okada, K., Sato, T., Goto, R., Sato, K., Ono, S., Schormann, T., Zilles, K., Fukuda, H. (2000). Correlation between human personality and neural activity in cerebral cortex. *NeuroImage, 11*, 541-546.

Suzuki, A., Fukasawa, T., Shiraishi, H., Ishii, G., Oshino, S., Aoshima, T., et al. (2007). No association between the TPH A218C polymorphism and personality traits in Japanese healthy subjects. *Progress in Neuropsychopharmacology & Biological Psychiatry, 31*(2), 395-398.

Svrakic, D., Przybeck, T., & Cloninger, C. (1992). Mood states and personality traits. *Journal of Affective Disorders, 24*, 217-226.

Svrakic, D. M., Draganic, S., Hill, K., Bayon, C., Przybeck, T. R., & Cloninger, C. R. (2002). Temperament, character, and personality disorders: etiologic, diagnostic, treatment issues. *Acta Psychiatrica Scandinavica, 106*(3), 189-195.

Svrakic, D. M., Whitehead, C., Przybeck, T. R., & Cloninger, C. R. (1993). Differential diagnosis of personality disorders by the seven-factor model of temperament and character. *Archives of General Psychiatry, 50*(12), 991-999.

Svrakic, N. M., Svrakic, D. M., & Cloninger, C. R. (1996). A general quantitative theory of personality development: Fundamentals of a self-organizing psychobiological complex. *Development and Psychopathology, 8*(1), 247-272.

Swelden, W. (1996). Wavelets: what next? *Proc IEEE, 84*(4), 680-685.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*: Thieme Medical Publishers, New York.

Tankard, C. F., Waldstein, S. R., Siegel, E. L., Holder, L. E., Lefkowitz, D., Anstett, F., et al. (2003). Cerebral blood flow and anxiety in older men: an analysis of resting anterior asymmetry and prefrontal regions. *Brain and Cognition, 52*(1), 70-78.

Ter-Pogossian, M. (1985). Three-dimensional biomedical imaging. In (Vol. 2, pp. 41-55). Florida: CRC Press.

Thuiller, W., Albert, C., Araúo, M. B., Berry, P. M., Cabeza, M., Guisan, A., et al. (2008). Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics, 9*(3-4), 137-152.

Turkheimer, F. E., Aston, J. A. D., Banati, R. B., Riddell, C., & Cunningham, V. J. (2003). A linear wavelet filter for parametric imaging with dynamic PET. *IEEE Transactions on Medical Imaging, 22*(3), 289-301.

Turkheimer, F. E., Brett, M., Aston, J. A. D., Leff, A. P., Sargent, P. A., Wise, R. J. S., et al. (2000). Statistical modeling of Positron Emission Tomography images in wavelet space. *Journal of Cerebral Blood Flow & Metabolism, 20*(11), 1610-1618.

Turner, R. M. (2004). Independent component analysis of personality and symptoms of depression and statistical parametric mapping of personality and brain function. Thesis. University of Canterbury, Christchurch, New Zealand.

Turner, R. M., Hudson, I. L., Butler, P. H., & Joyce, P. R. (2003). Brain function and personality in normal males: A SPECT study using statistical parametric mapping. *NeuroImage, 19*(3), 1145-1162.

Üzümcü, M., Frangi, A. F., Reiber, J. H. C., & Lelieveldt, B. P. F. (2003). Independent Component Analysis in Statistical Shape Models. SPIE Medical Imaging - The International Society for Optical Engineering, 5032, 375-383.

Vaidya, J. G., Gray, E. K., Haig, J., & Watson, D. (2002). On the temporal stability of personality: Evidence for differential stability and the role of life experiences. *Journal of Personality and Social Psychology, 83*(6), 1469-1484.

Van de Ven, V. G., Formisano, E., Prvulovic, D., Roeder, C. H., & Linden, D. E. J. (2004). Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. *Human Brain Mapping, 22*(3), 165-178.

Van Hateren, J. H., & Ruderman, D. L. (1998b). Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences, 265*(1412), 2315-2320.

Van Hateren, J. H., & Van der Schaaf, A. (1998a). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences, 265*(1394), 359-366.

Vidakovic, B. (1998). Nonlinear Wavelet Shrinkage with Bayes Rules and Bayes Factors. *Journal of the American Statistical Association, 93*(441), 173-179.

Vidakovic, B. (1999). *Statistical modelling by wavelets*. New York: John Wiley & Sons.

Videbech, P., Ravnkilde, B., Pedersen, A., Egander, A., Landbo, B., Rasmussen, N., et al. (2001). The Danish PET/depression project: PET findings in patients with major depression. *Psychological Medicine, 31*(7), 1147-1158.

Vig'ario, R., Sarela, J., Jousmiki, V., Hamalainen, M., & Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering, 47*(5), 589-593.

Vigneron, V., Paraschiv-Ionescu, A., Azancot, A., Sibony, O., & Jutten, C. (2003). *Fetal electrocardiogram extraction based on non-stationary ICA and wavelet denoising. 7th* Symposium on Signal Processing and Appl., Paris, France, 69–72.

Visser, M. E., & Both, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B: Biological Sciences, 272*(1581), 2561-2569.

Walker, J. S. (1999). A Primer on Wavelets and their Scientific Applications. Boca Raton: CRC Press, .

Walker, J. S. (2000). Combined image compressor and denoiser based on tree-adapted wavelet shrinkage. *Optical Engineering,, 41*, 1520-1527.

Walker, J. S. (2002). Tree-Adapted Wavelet Shrinkage. Advances in Imaging and Electron Physics. (pp. 343-394). New York: Elsevier Science (USA).

Walker, J. S. (2003). *New methods in wavelet-based image denoising." Progress in Analysis".* Paper presented at the International ISAAC Congress, August, New Jersey, USA.

Walker, J. S. (2008). *FAWAVE (A Fourier / Wavelet Analyzer)*, from http://www.uwec.edu/walkerjs/FAWAVE

Walker, J. S., & Chen, Y. J. (2000). Image denoising using tree-based wavelet subband correlations and shrinkage. *Optical Engineering, , 39*, 2900-2908.

Walnut, D. F. (2004). An Introduction to Wavelet Analysis: Birkhauser.

Walther, G. R., Hughes, L., Vitousek, P., & Stenseth, N. C. (2005). Consensus on climate change. *Trends in Ecology & Evolution, 20*(12), 648-649.

Wand, M. P., & Jones, M. C. (1995). *Kernel smoothing*: Chapman & Hall/CRC.\

Wang, L. (2005). Change in non-parametric regression with long memory errors. *Statistics & Decisions, 23*(2), 147-159.

Wang, L. (2008). Change-Point Estimation in Long Memory Nonparametric Models with Applications. *Communications in Statistics-Simulation and Computation, 37*(1), 48-61.

Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika, 82*(2), 385-397.

Wang, Y. (1996). Function estimation via wavelet shrinkage for long-memory data. *The Annals of Statistics, 24*(2), 466-484.

Wang, Y. (1998). Change Curve Estimation Via Wavelets. *Journal of the American Statistical Association, 93*(441), 163-172.

Weaver, J. B., Yansun, X., Healy, D. M. J., & Cromwell, L. D. (1991). Filtering noise from images with wavelet transforms. *Magnetic Resonance in Medicine, 24*, 288-295.

Whitcher, B. (2005a). Wavelet Analysis. In B. S. Everett & C. Palmer (Eds.), *Encyclopaedic Companion to Medical Statistics* (pp. 367-368): Hodder Arnold.

Whitcher, B. (2005b). Statistics in Imaging. In B. S. Everett & C. Palmer (Eds.), *Encyclopaedic Companion to Medical Statistics* (pp. 333-334): Hodder Arnold.

Whitcher, B., Guttorp, P., & Percival, D. B. (2000). Wavelet Analysis of Covariance with Application to Atmospheric Time Series. *Journal of Geophysical Research, 105*(D11), 14,941-962.

Whitcher, B., & Jensen, M. J. (2000). Wavelet estimation of a local long memory parameter. *Exploration Geophysics, 31*(1), 94-103.

Whitcher, B., Schwarz, A. J., Barjat, H., Smart, S. C., Grundy, R. I., & James, M. F. (2005). Wavelet-based cluster analysis: data-driven grouping of voxel time courses with application to perfusion-weighted and pharmacological MRI of the rat brain. *NeuroImage, 24*(2), 281-295.

Whitcher, B., Whitcher, M. B., & Gpl, L. (2006). The waveslim Package. *exchange, 20*.

White, M. A., Brunsell, N., & Schwartz, M. D. (2003). Vegetation phenology in global change studies. In M. D. Schwartz (Ed.), *Phenology: An integrative environmental science*,453-466.

Wiener, N. (1964). Extrapolation, interpolation, and smoothing of stationary time series. New York, USA: Wiley.

Wilcox, R. R. (2005). Introduction to robust estimation and hypothesis testing: Academic Press.

Wink, A. M., Bernard, F., Salvador, R., Bullmore, E., & Suckling, J. (2006). Age and cholinergic effects on hemodynamics and functional coherence of human hippocampus. *Neurobiology of Aging, 27*(10), 1395-1404.

Wishart, J. (2009). Kink estimation with correlated noise. *Journal of the Korean Statistical Society, 38*(2), 131-143.

Wishart, J., & Kulik, R. (2010). Kink estimation in stochastic regression with dependent errors and predictors. *Arxiv preprint arXiv:1003.1535*.

Witkin, H. A., Lewis, H. B., Hertzman, M., Machover, K., Meissner, P. B., & Wapner, S. (1977). *Personality through perception*: Greenwood Press Conn.

Wornell, G. W. (1993). Wavelet-based representations for the 1/f family offractal processes. *Proceedings of the IEEE, 81*(10), 1428-1450.

Wornell, G. W., & Gaumond, C. F. (1999). Signal processing with fractals: a wavelet based approach. Prentice Hall.

Worsley, K. (1995b). Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *Annals of Statistics, 23*,640-669.

Worsley, K., Marrett, S., Neelin, P., Vandal, A., Friston, K., & Evans, A. (1995a). A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping, 4*, 58-73.

Worsley, K. J. (2003). Developments in random field theory. *Human Brain Function*, 881?886.

Worsley, K. J., Evans, A. C., Marrett, S., & Neelin, P. (1992). A three-dimensional statistical analysis for CBF activation studies in human brain. *Journal of Cerebral Blood Flow and Metabolism, 12*, 900-918.

Xing, C. J., & Wang, S. J. (2001). A New Filtering Algorithm Based on Extremum and Median Value. *Journal of Image and Graphics, 6A*(5), 533-536.

Yang, R., Yin, L., Gabbouj, M., Astola, J., & Neuvo, Y. (1995). Optimal weighted median filters under structural constraints. *Signal Processing, 43*, 591--604.

Ylipaavalniemi, J., Mattila, S., Tarkiainen, A., & Vigario, R. (2006). Brains and phantoms: an ICA study of fMRI. *Lecture Notes in Computer Science, 3889*, 503-510.

Zald, D. H., Curtis, C., Folley, B. S., & Pardo, J. V. (2002). Prefrontal contributions to delayed spatial and object alternation: a positron emission tomography study. *Nueropsychology-New York, 16*(2), 182-189.

Zaroubi, S., & Goelman, G. (2000). Complex denoising of MR data via wavelet analysis: application for functional MRI. *Magnetic Resonance Imaging, 18*(1), 59-68.

Zeileis, A., Kleiber, C., Kräer, W., & Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics and Data Analysis, 44*(1-2), 109-123.

Zeileis, A., Leisch, F., Hornik, K., & Kleiber, C. (2002). strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software, 7*(2), 1-38.

Zhang, H., & Singer, B. (1998). *Recursive Partitioning in the Health Sciences.* New York: Springer-Verlag.

Zuckerman, M., & Stelmack, R. M. (2004). On the Psychobiology of Personality: Essays in Honor of Marvin Zuckerman: Pergamon.